An ELECTRONICS TODAY INTERNATIONAL publication

eti
CIRCUIT
techniques
volume one

The how, what, which, where,
why and how much anthology
of electronic components,
circuits and techniques.

ETI
Circuit
Techniques
volume one

Editor: Jan Vernon
Layout: Bill Crump
Managing Editor: Collyn Rivers

CONTENTS

OP-AMPS

An operational amplifier is just a high gain amplifier — you stick a voltage in and a much larger one comes out. But you'd never know this from the data sheets. 'Overkill' confuses all but the most experienced. It really doesn't have to be so. Tim Orr explains. . .

OP-AMPS HAVE TWO inputs, inverting and non-inverting, denoted by − and + respectively. The op-amp amplifies the difference in the voltages applied to these two inputs, the output going positive if the + input is positive with respect to the − input, and vice versa. Without extra circuitry, though, an op-amp is virtually useless, for the gain is too high to be useable and distortion is excessive. Fortunately both parameters can be controlled by feedback.

An op-amp with negative feedback is shown in Fig. 1. Two resistors set the closed loop voltage gain, and as long as this is small compared to the open loop gain, it will be determined by the resistor ratio RF/RI. The open loop gain, the voltage gain when RF is removed, is typically 1 000 000. This massive gain is clearly much too large to be used without feedback. Closed loop voltage gains of 100 are about as much as it is practical to use.

Biased example

The arrangement in Fig. 1 is known as a 'virtual earth' amplifier. The non-inverting input is connected to earth, and the inverting input is maintained by the feedback applied via RF at a voltage which is virtually earth potential.

The input impedance of the amplifier in Fig 1 is simply R1. The output impedance is a little more complicated, approximately: −

$$\frac{\text{output impedance of the op amp} \times \text{closed loop gain}}{\text{open loop gain}}$$

Suppose we want an amplifier with a gain of 10, and an input impedance of 1M. This means that RI is 1M. Therefore RF must be 10 M (see Fig. 2). With a 1 V sinewave as the input signal we get a 10 V sinewave as the output. However, when the input signal is held at 0V, it is positive! This is an error

voltage, which may be undesirable. The cause of the problem is the 'INPUT BIAS CURRENT' of the op-amp. The input of many op-amps looks like the circuit shown in Fig. 3. If these transistors are to operate correctly they need a standing emitter current which implies that they need an input base current. It is this base current which is the op-amp's 'INPUT BIAS CURRENT'. For a 741 this current can be as large as 0.5 μA. In the arrangement of Fig. 2 this current can only come through RF, which means that, the output voltage could be as large as 0.5 μA x 10 M, which is +5 V! One way to remedy this error is to use the circuit shown in Fig. 4. A resistor has been inserted between the non-inverting input and ground. This resistor has the value of RF in parallel with RI. It allows both the inputs to sink slightly and thus maintain the voltage balance at the inputs. The output voltage is then nearly 0 V. However, the two input transistors may not be that well matched, so the bias currents into each input may be different. This is known as the 'INPUT OFFSET CURRENT' and its effect can be nulled by making the 910 k resistor in Fig. 4 a variable resistor. But even if the bias currents (for say a 741) were zero, then the output voltage would still not be 0 V.

Get set, they're off

The output voltage could range between ±60 mV. This is due to the 'INPUT OFFSET VOLTAGE' which for a 741 can be as much as ±6 mV, which is then multiplied by the closed loop voltage gain of the stage (in this case 10 giving us ±60 mV. This can be compensated by using the circuit shown in Fig. 5. Terminals 1 and 5 on a 741 can be used to compensate for the input offset voltage. The input offset voltage is the V_{be} imbalance between the two input transistors.

Now that we know how to eliminate the spurious dc offsets, we can try designing some dynamic circuits and find out why they don't work as expected! For example, try

CLOSED LOOP VOLTAGE GAIN $= -\frac{RF}{RI} = \frac{V_{OUT}}{V_{IN}}$ (AV)
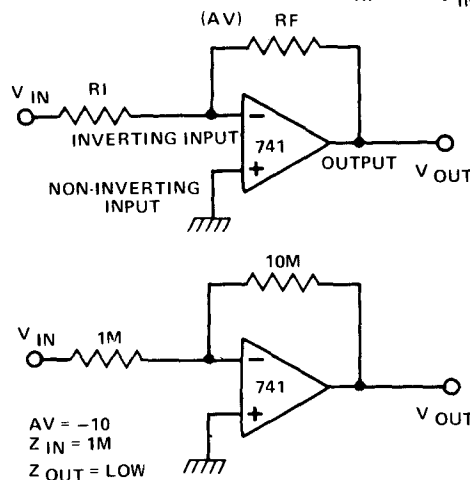
Figs 1 and 2 show (upper) the basic inverting op-amp stage. Gain is given by the ratio of resistors RF/RI, input impedance is simply RI, while the output impedance is more complicated (see text). Fig. 2 (lower) shows a stage with a gain of 10 and an input impedance of 1 M.

AV = −10
Z IN = 1M
Z OUT = LOW

NPN DIFFERENTIAL PAIR
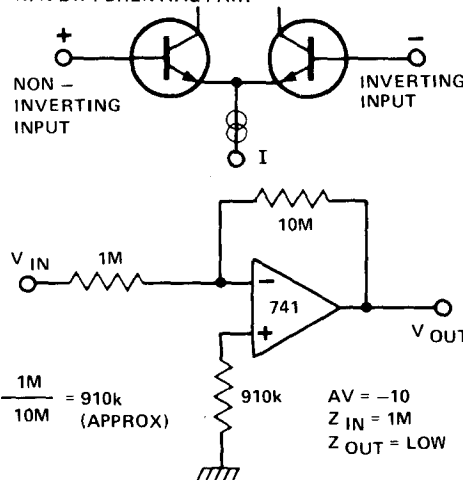
Figs 3 and 4 (left) show (upper) a typical op-amp input stage. This is a differential amplifier made up of a pair of NPN transistors driven by a constant current source. Fig.4 (left) shows a 910 k resistor in series with the + input of the op-amp. This reduces the effects of the input offset current.

$\frac{1M}{10M} = 910k$ (APPROX)

AV = −10
Z IN = 1M
Z OUT = LOW

putting a 1 V sinewave at 200 kHz into the circuit shown in Fig. 5. What you would expect is a 10 V, 200 kHz sinewave at the output — but you don't get one. What appears is a rather bent 200 kHz triangle waveform. This is because the 'SLEW RATE' of the op-amp has been exceeded. The slew rate is the speed at which the output voltage can move, and for a 741 is typically 0.5 V/μs when it crosses zero, so the op-amp, faced with this demand, just gives up and slew limits, drawing out straight lines as it does so.

Listen to the band(width)

Another limitation is 'BANDWIDTH'. A 741 has a GAIN BANDWIDTH product of approximately 1 MHz. This means that the product of the voltage gain times the operating frequency cannot exceed 1 MHz.

For example, if you want the amplifier to have a gain of 100, then the maximum frequency at which this gain can be obtained is 10 kHz. Figure 6 illustrates this phenomenon. Curve A is the open loop response, note that the voltage gain is 1 at 1 MHz, hence the gain bandwidth product of 1 MHz. The slope of the curve is −20 dB/decade, which is caused by a single 30 pf capacitor inside the IC. Now, if the resistor ratio is set to give a voltage gain of 100, then the op-amp gives a frequency response shown by curve C, which is flat up until 10 kHz. A gain off 10 rolls off at 100 kHz (D) and a gain of 1 000 rolls off at 1 kHz (B). Thus it is very easy to see just what the closed loop frequency response will be. However, don't forget the slew rate problem. You may be able to construct an amplifier with a voltage gain of 10, which works up to 100 kHz, but the output voltage will be limited to less than 3 Vpp! Another problem is distortion in the op-amp. Negative feedback is used to iron out any distortion generated by the op-amp, but negative feedback relies on there being some spare voltage gain available. For instance, say the op-amp generates 10% distortion and there is a surplus voltage gain of 1 000,

i.e. $\left(\dfrac{\text{open loop gain}}{\text{closed loop gain}} \right)$,

then the distortion will be reduced to approximately,

$$\dfrac{\text{open loop distortion}}{\text{surplus voltage gain}} = \dfrac{10\%}{1\ 000} = 0.01\%$$

So, negative feedback is used to eliminate distortion products, However, if there is no surplus voltage gain, as in the case of a 741 amplifier working at 10 kHz, with a closed loop gain of 100, the distortion will rise dramatically at this point.

Current thinking

Most op-amps have a voltage output, although some have a current output. If you short-circuit a voltage output then large currents could flow and thermal destruction might follow. To overcome this problem, most op-amps have a current limited output so that they can tolerate an indefinite short to ground. A 741 is limited to about 25 mA. Another current of note is the supply 'BIAS CURRENT'. This is the current consumed when the op-amp is not driving any load. For a 741 this current is typically 2 mA, which makes it unsuitable for some battery applications.

There are some op-amps which can be programmed by inserting a current into them so that their supply current can be controlled. This means that they consume only micropower when in their 'standby' mode, and can be quickly turned on to perform a particular task.

Voltages differently

In the few examples shown so far, the op-amp has been used to amplify voltages which have been generated with respect to ground. However, sometimes, it is required to measure the difference between two voltages. In this case you would use a 'Differential' amplifier, Fig. 7. By using two matched pairs of resistors, the formula for the voltage gain is made very simple. It is thus possible to sumperimpose a 1 V sinewave on both the inputs, and yet have the output of the amplifier ignore this common mode signal and amplify only differential signals. The amount by which the common mode signal is rejected is called the CMRR (the Common Mode Rejection Ratio) and is typically 90 dB for a 741. Thus a common mode 1 V signal would be reduced to 33 μV.

Another rejection parameter to be noted is the supply voltage rejection ratio. For a 741 the typical rejection is 90 dB; that is, if the power supply changes by 1 V the change in the output voltage will be 33 μV.

When designing with op-amps it is very important to know what voltage range the inputs will work over, and the maximum voltage excursion you can expect at the output. For instance, the 741 can operate with its inputs a few volts from either power supply rail, and its inputs can withstand a differential voltage of 30 V (with a power supply of 36 V).

This is not true of all op-amps, some have a very limited differential input voltage range, for instance the CA3080 will zener when this voltage exceeds 5 V and the amplifier performance will then be drastically changed.

The output excursion of the op-amp is also important. The 741 can only typically swing within about 2 V of either supply rail, whereas the CMOS op-amp can swing to within 10mV of either rail so long as the load into which they are driving is a very high impedance.
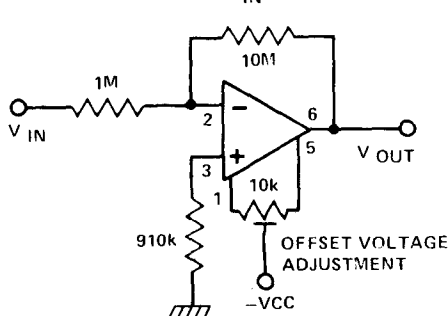
VOLTAGE GAIN $A_V = \dfrac{V_{OUT}}{V_{IN}}$

Fig. 5. A variable resistor connected between pins 1 and 5 of a 741 can be used to reduce the effects of the input offset voltage.
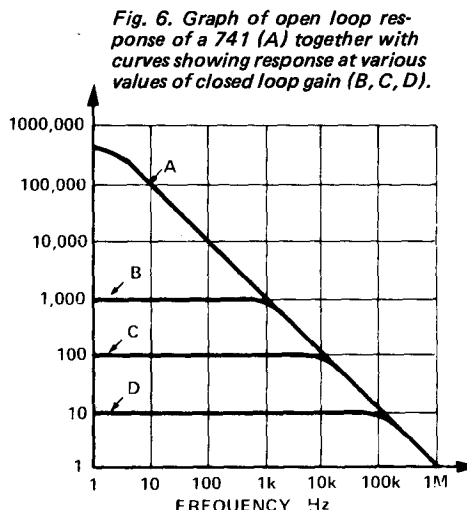
Fig. 6. Graph of open loop response of a 741 (A) together with curves showing response at various values of closed loop gain (B, C, D).
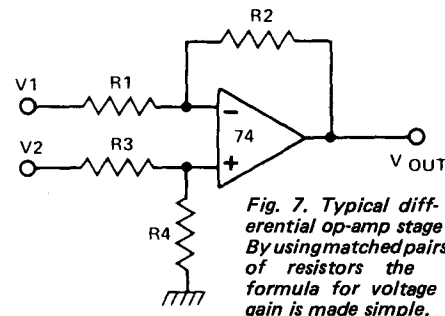
Fig. 7. Typical differential op-amp stage By using matched pairs of resistors the formula for voltage gain is made simple.

$$V_{OUT} = \left[\left(\dfrac{R1 + R2}{R3 + R4} \right) \dfrac{R4}{R1}\ V2 \right] - \dfrac{R2}{R1}\ V1$$

BUT IF WE MAKE R1 = R3
AND R2 = R4

THEN $V_{OUT} = \dfrac{R2}{R1}\ (V2 - V1)$

12 V REGULATED POWER SUPPLY

The large open loop voltage gain of an op-amp is very useful in providing a regulated low output impedance power supply. A 5V1 voltage reference is generated by a zener diode ZD1 (this voltage reference could be made more stable by running it at constant current). A PNP transistor is used as a series regulator. However, this transistor inverts the signal from the op-amp output, and so, in order to get negative feedback, the feedback is taken to the non-inverting input! The operations is as follows. The inverting input is held at 5V1. If the 'PSU OUTPUT' tries to fall, the voltage at the non-inverting input falls. Therefore the op-amp's output will also fall, thus turning on the PNP transistor which then pulls up the 'PSU OUTPUT'. Thus the output voltage is stabilised. Also, the output impedance is very low, due to this negative feedback. The output impedance at high frequencies (where the op-amp gain is low) is further reduced by the 10 μ capacitor. To squeeze the last drop of voltage out of the system, before a collapsing unregulated supply rail causes the regulated supply to drop out, a 5V1 zener diode (ZD2) has been included. This allows the op-amp output to work at about 7 volts below the unregulated supply rail. Thus, a regulated output is maintained until the PNP transistor saturates. This means that the unregulated rail can fall to within about 200 mV of the regulated rail!

+VE UNREGULATED SUPPLY 30V→12.2V
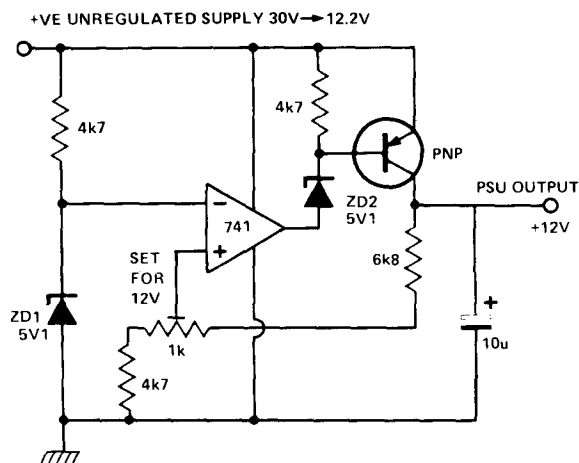
SIMPLE INTEGRATOR

An op-amp and a capacitor can be used to implement, to a high degree of accuracy, the mathematical process of integration. In this case, current is summed over a period of time and the resultant voltage generated is the integral of that current as a function of time. What this means that if a constant voltage is inputted to the circuit, a ramp with a constant slope is generated at the output. When the input is positive, the output of the op-amp ramps negative.

In doing so it pulls the inverting terminal negative so as to maintain a 'virtual earth' condition. In fact the input current (Vin/R1) is being equalled by the current flowing through the capacitor, thus equilibrium is maintained. The equation governing the behaviour of a capacitor is C x dV/dt = i, where dV/dt is the rate of change of voltage across the capacitor.
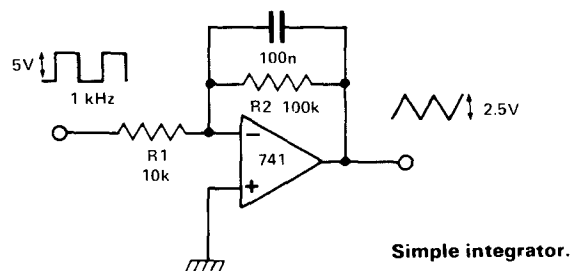
Therefore

$$\frac{dV}{dt} = \frac{i}{C}$$

Thus

$$\frac{dV}{dt} = \frac{Vin}{R1C}$$

So, when a square wave is applied to the circuit in Fig. 10, triangle waveforms are generated. R2 was added to provide DC stability. Its inclusion does slightly corrupt the

Simple integrator.

mathematical processes, but not enormously. A good point about this integrator design is that it has a very low output impedance. You can put a load on the output and the op-amp will still generate the same waveform — that's what is so nice about negative feedback.
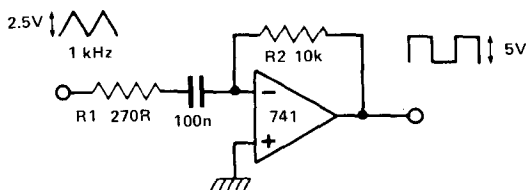

Fig. 11. Simple differentiator.

SIMPLE DIFFERENTIATOR

Mathematically, differentiation is the reverse process to integration. Thus, in the differentiator circuit the C and the R are reversed with respect to the integrator circuit.

The input waveform is a triangle with a constant rise and fall slope. This constant slope, when presented to a capacitor will generate a constant current. When the slope direction reverses, then so will the current flow. This current when passed through a resistor (R1), will then generate a square wave.

TRIANGLE SQUARE OSCILLATOR

A Schmitt trigger and an integrator can be used to construct a very reliable oscillator which generates triangle and square waveforms. The operation of the circuit is very simple and always self starting. The Schmitt trigger is formed from IC1, the integrator from IC2. Suppose the output of the Schmitt is positive. This will cause the integrator to generate a negative going ramp. This ramp is then fed back to the input of the Schmitt. When the lower hysteresis level has been reached the output of the Schmitt snaps into its negative state, current is taken out of the integrator which then generates a positive going ramp. The integrator's output ramps up and down between the upper an lower hysteresis levels. The speed at which the integrator moves is determined by the magnitude of the voltage applied to it. In this circuit, the magnitude of the voltage and hence the oscillation frequency, are controlled by a potentiometer, giving a 100 to 1 control range. This
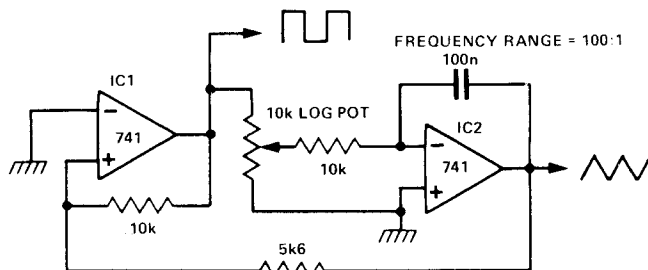
circuit is the basis of most function generators. By bending the triangle it is possible to synthesise an approximation to a sinewave. With a bit more electronics it is also possible to make the oscillator voltage controlled.
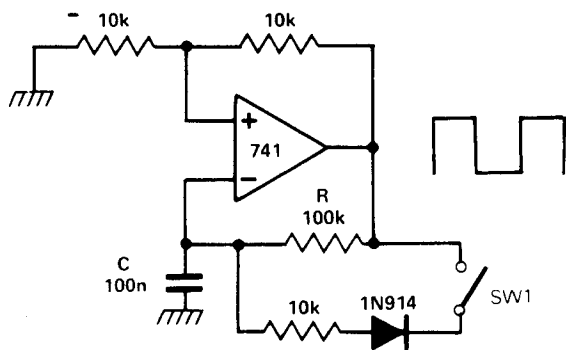
SINGLE OP-AMP OSCILLATOR

This circuit has a Schmitt trigger and a 'sort of integrator' all built around one op-amp. The positive feedback is via the 10 k resistors. The 'integration', (the timing) is controlled by the RC network. The voltage at the inverting input follows that of the RC charging exponential, except that it is confined to be within the upper and lower hysteresis levels. Thus the hysteresis levels and the RC time constant determine the frequency of operation. It is possible to make the output square wave have a large mark to space ratio. By closing the switch SW1, the discharge time of the capacitor becomes eleven times faster than the rise time. Thus a square wave with an 11:1 mark space ratio is generated.
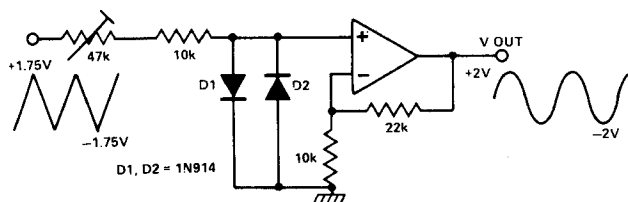
SIMPLE TRIANGLE TO SINEWAVE CONVERTER

Here is a simple way of converting a triangle to a sinewave. The logarithmic characteristic of the diodes is used to approximate that of a sine curve. Distortion is 5% or so. However, the distortion may be tolerable if the sinewave is only used to generate audio tones.

LINEAR VOLTAGE CONTROLLED OSCILLATOR

This oscillator is very similar to the triangle square wave oscillator shown on this page, except that this one is voltage controlled. The integrator and Schmitt trigger action are the same as before, but the feedback has been altered. The input voltage Vin, is applied differentially to the integrator via the resistor network. The larger the value of Vin, the faster the integrator ramps up and down. Thus the frequency of the operation is determined by an external positive control voltage. The frequency is linearly proportioned to this control voltage.

When the output of the Schmitt is low, Q1 is off and all the input voltage is applied to the inverting input. Half of the input voltage is always applied to the non-inverting input. Therefore the integrator's output ramps downward until the Schmitt flips into its positive state. Now, Q1 is switched on and the voltage at the inverting input is negative with respect to the non-inverting input. Hence the integrator now ramps upwards.
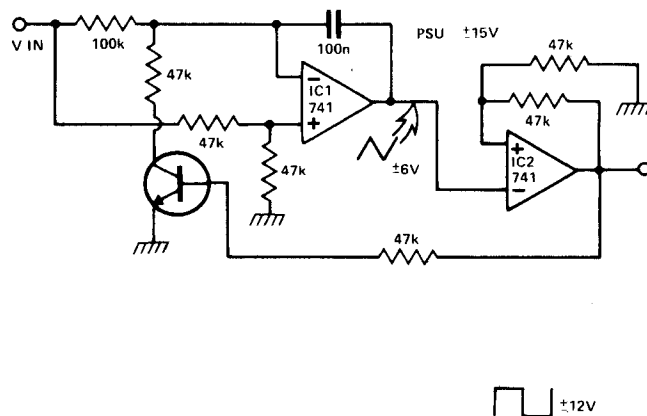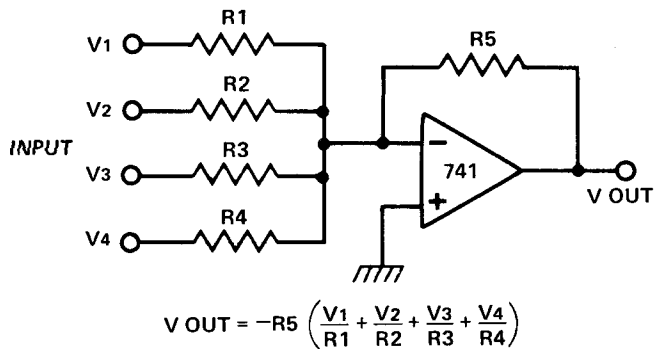
BASIC SUMMING CIRCUIT (MIXER)

A virtual earth amplifier can be used to mix several signals together. The output voltage is a mixture of all the inputs. The amount of an input that appears at the output is inversly proportional to the input resistor. If the input voltages are fed into potentiometers before being fed to the mixer, then their individual levels can be manually adjusted. This is the basis of most audio mixers, although the cheaper units use op-amps. Most op-amp mixers will degrade the signal to noise ratio of the signals by more than a good discrete component amplifier.

$$V\ OUT = -R5 \left(\frac{V1}{R1} + \frac{V2}{R2} + \frac{V3}{R3} + \frac{V4}{R4} \right)$$

TURNING A LINEAR POT INTO A LOG POT

By using the virtual earth characteristic of an op amp, a linear pot can be made to have the characteristics of a log pot. It seems to be fair to say that low cost linear pots are far more linear than log pots are logarithmic. Thus the linear pot can be turned into a better log pot than the actual log pot itself. By varying the resistor ratio 5k6 to 50 k, other laws can be produced, such as something in between log and linear or maybe a law that is even more extreme than log.
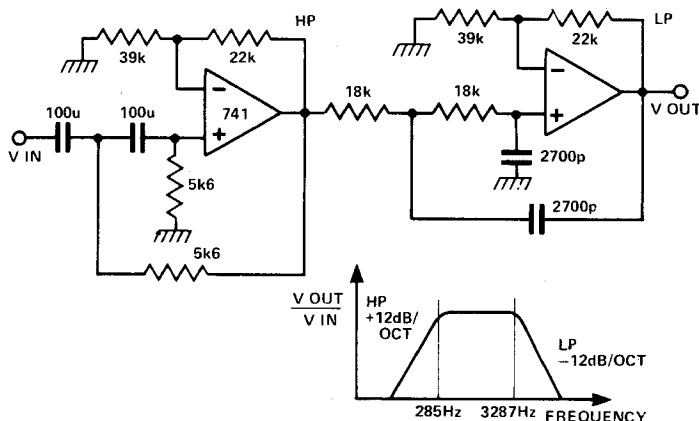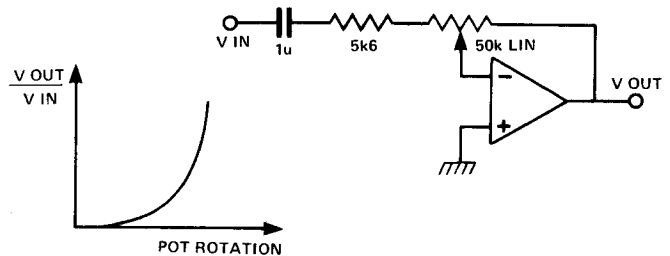
SIMPLE SPEECH FILTER

The telephone system has been designed for speech communication. The bandwidth of the system is 300 Hz to 3400 Hz, which has been arrived at after many years of experimentation. Thus, it is true to say that much of the information in speech is contained between these frequency limits. The circuit shows a filter structure that will simulate the telephone bandwidth. It could have many uses, for instance as a 'speech filter' for noisy radio reception or land line communications, or as a voice detector for a light show.

SIMPLE MUSICAL CHIME GENERATOR

The circuit shown is that of a multiple feedback band-pass filter. The present is used to add some positive feedback and so further increase the Q factor. The principle of operation is as follows. A short click (pulse) is applied to the filter and this makes it ring with a frequency which is its natural resonance frequency. The oscillations die away exponentially with respect to time and in doing so closely resemble many naturally occuring percussive or plucked sounds. The higher the Q the longer the decay time constant. High frequency resonances resemble chimes, whereas lower frequencies sould like claves or bongos, By arranging several of these circuits, all with different tuning, to be driven by pulses from a rhythm generator an interesting pattern of sounds can be produced. There may be some stability problems when high Q or high frequency operation is involved. To achieve better performance, an op-amp with a greater bandwidth than the 741 should be used.

Alternatively, a different structure, such as a state variable filter could be used. Qs of up to 500 can be obtained with this latter circuit.

$$f_o = \frac{1}{185 \times 1000 \times C}\ Hz$$
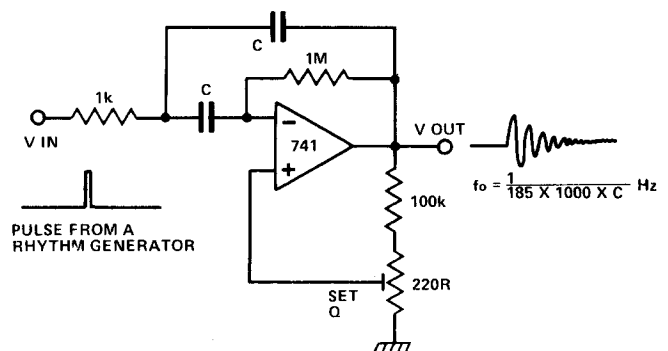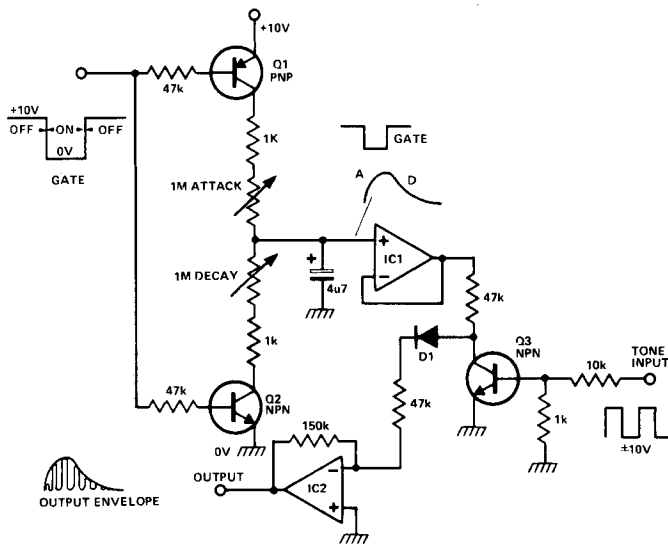
MUSICAL ENVELOPE GENERATOR AND MODULATOR

A gate voltage is applied to initiate the proceedings. When the gate voltage is in the ON state, Q1 is turned on, and so the capacitor C is charged up via the attack pot in series with the 1 k resistor. By varying this pot, the attack time constant can be manipulated. A fast attack gives a percussive sound, a slow attack the effect of 'backward' sounds. When the gate voltage returns to its off state, Q2 is turned on and the capacitor is then discharged via the decay pot and the other 1 k resistor, to ground. Thius the decay time constant of the envelope is also variable.

This envelope is buffered by IC1, a high impedance voltage follower and applied to Q3 which is being used as a transistor chopper. A musical tone in the form of a squarewave is connected to the base of Q3. This turns the transistor on or off and thus the envelope is chopped up at regular intervals, the intervals being determined by the pitch of the squarewave.

The resultant waveform has the amplitude of the envelope and the harmonic structure of the squarewave. IC2 is used as a virtual earth amplifier to buffer the signal and D1 ensures that the envelope dies away at the end of a note.

## SIMPLE MUSICAL ENVELOPE GENERATOR

A simple generator can be constructed using the CA3080 (made by RCA). This circuit will also enable the use of an audio waveform the harmonic structure of which will not be significantly affected as it is modulated. The CA3080 is an op amp with a difference. It has a current output and an extra input into which a current, I_A is fed. The output is the product of the input voltage X I_A. Thus the I_A can be used to control the amplifier's gain.

The input voltage range for low distortion operation is very low, of the order of ± 25 mV.

The CA3080 is being used as a two-quadrant multiplier. A small voltage, (± 25 mV), is applied to its non-inverting input. When the switch S1 is closed, the capacitor C is charged up and a current of about 150 μA flows into the IA input terminal. When S1 is opened, C discharges through the 150 k resistor into the I_A input. This current dies away exponentially. As the output is the product of the input voltage X I_A, then
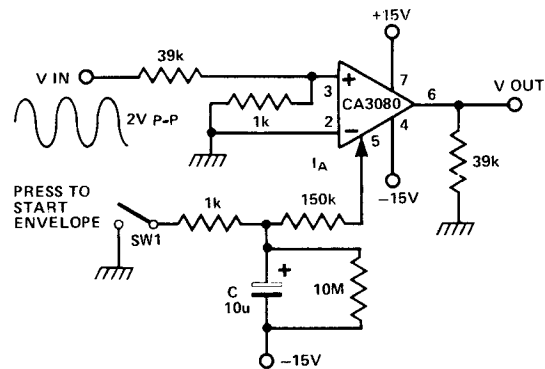
an exponential envelope is generated. Breakthrough after the decay is very good, better than −80 dB.
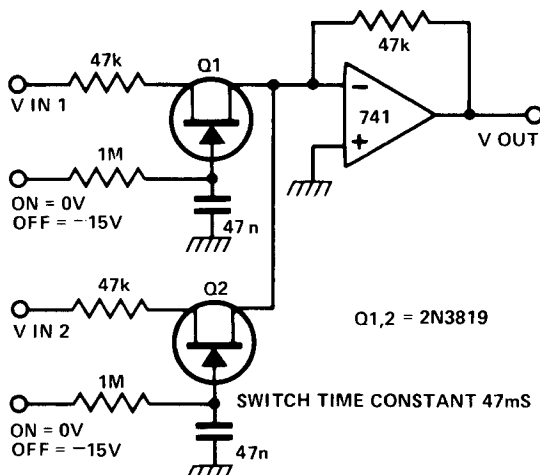
SILENT AUDIO SWITCHING

Sometimes electronic switches for audio signals are required. FETs can be used to perform the switching, but they can cause distortion, the resultant output impedance is not very low and clicks generated by the switching signal can break through. The circuit shown virtually eliminates all of these problems. By using an op-amp a very low output impedance is obtained as well as the possibility of selecting or mixing one or more of many input channels. Because of the virtual earth mixing, the voltage across any FET that is switched on is very small. If the output voltage is 1V and the FETs ON resistance is 470R, then the voltage across the FET is about 10 mV. When large voltages are applied to a turned on FET, the distortion is large, but if the voltage is small, (10 mV say), the distortion could be less than 0.1%. Thus the virtual earth mixing enables low distortion operation. Lastly, to stop the generation of switching clicks, a time constant of 47 msec has been enforced at the gate of the FETs.

LED BAR PPM DISPLAY FOR AUDIO

The peak voltage detector can be used to control an illuminated audio level monitor having the same characteristics as a PPM (Peak Programme Meter). A bar column of LEDs is arranged so that as the audio signal level increases, more LEDs in the column light up. The LEDs are arranged vertically in 6 dB steps. A fast response time and a one second decay time has been chosen so as to give an accurate response to transients and a low 'flicker' decay characteristic. The op-amps that drive the LEDs are used as comparators. On each of their inverting inputs they have a dc reference voltage, which increases in 6 dB steps up the chain. All of their non-inverting inputs are tied together and connected to the positive peak envelope of the audio signal. Thus as this envelope exceeds a particular voltage reference, that op-amp output goes high and the LED lights up. Also, all the LEDs below this are illuminated.
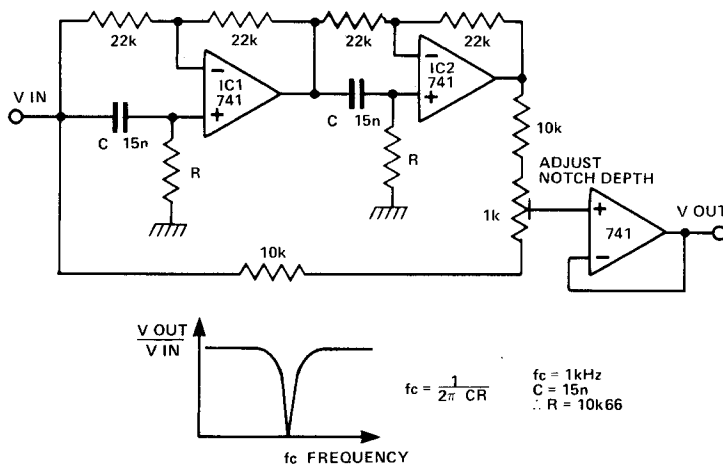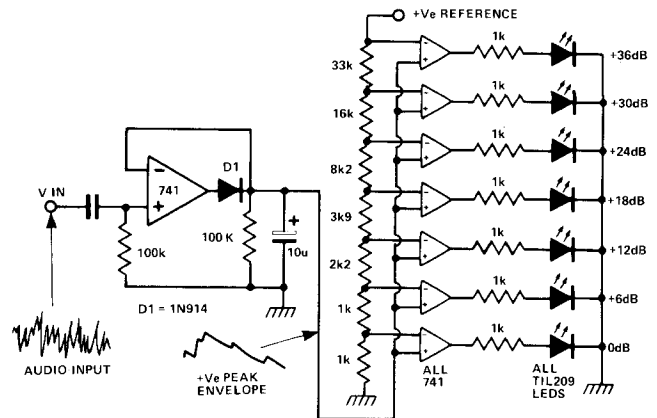

ALL-PASS NOTCH FILTER

Sometimes when processing analogue signals a constant tone causes a nuisance and so an active filter is called upon to 'notch' it out. The filter can be tuned so that its notch is at exactly the same frequency as the unwanted signal so that it can be selectively attenuated. This method is sometimes used to remove mains hum. The circuit works as follows IC1 and 2 are a pair of all-pass filters. These filters have a flat frequency response, but their phase changes with frequency. Their overall maximum phase shift is 360°, a phase shift of 180° occurring at a frequency of 1/2CR Hz. At this frequency the signals are inverted. Thus, by mixing the phase delayed signal with the original, cancellation can be produced which forms a notch in the frequency response. The preset is used to get the deepest notch available. The operating frequency can be changed by varying the two resistors R. For instance for 50 Hz operation, R should be:—

$$10.66k \times \frac{1000}{50} = 213.2k \quad \text{Nearest E12 fit is 220k}$$

$$fc = \frac{1}{2\pi \, CR}$$

fc = 1kHz
C = 15n
∴ R = 10k66

NOISE GENERATOR

The zener breakdown of a transistor junction is used in many circuits as a noise generator. The breakdown mechanism is random and so generates a small noise voltage. Also this voltage has a high source impedance. By using the op-amp as a high input impedance, high ac gain amplifier, a low impedance, large signal noise source is obtained. The preset is used to set the noise level by varying the gain from 40 to 20 dB.
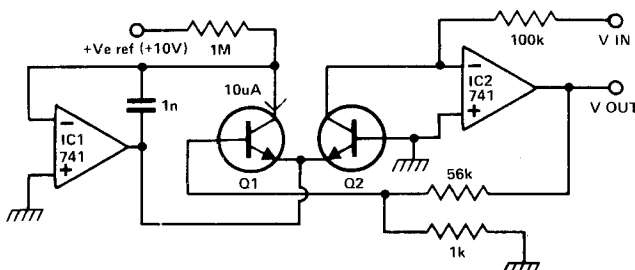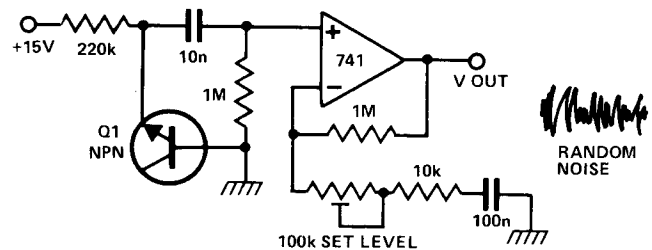
## LOGARITHMIC VOLTAGE TO VOLTAGE CONVERTER

The output voltage is logarithmically proportional to the input voltage. The difference between this circuit and the previous is that the exponentiator is in the feedback loop of the op-amp and hence the mathematical function has been inverted. The circuit is useful for performing true logarithmic compression or for converting linear inputs into dBs.

10

EXPONENTIAL VOLTAGE TO CURRENT/VOLTAGE CONVERTER

The circuit shown converts a linear input voltage into an exponential current or voltage. This type of circuit is used in music synthesisers to change linear control voltages into musical intervals. That is, if the circuit were used to control an oscillator, input increments of 1 V would change the pitch by one octave. The exponential characteristics of a transistor are exployed to generate the correct transfer function. Q1 and !2 are matched pairs of transistors, preferably a transistor dual. IC1 maintains Q1 at a constant current. Thus, the op-amp serves only to bias the emitter of the second transistor Q2 into a suitable operating region. The purpose of Q1 is to generate this bias voltage. The base emitter junction of a transistor
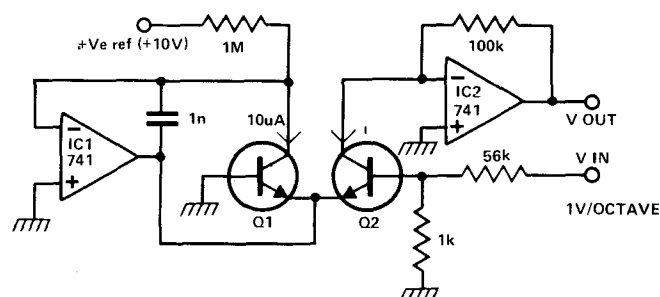
has a high temperature coefficient (−1.9 mV/°C) and so the reason for using a matched pair is to use the first transistor, Q1, to provide temperature compensation for the second.
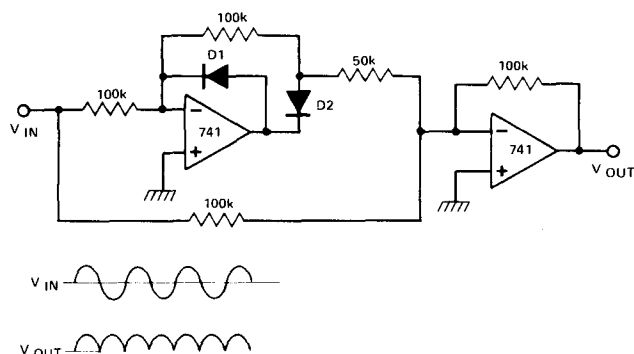
PRECISION HALF WAVE RECTIFIER

Rectifying small signals with any accuracy can be very difficult using diodes only due to their forward voltage drop of about 0.6 V. However, an op-amp can be used to reduce this voltage drop to virtually nothing. Consider the circuit shown. There is negative feedback so that 'virtual earth' circumstances exist. When Vin is positive, D1 conducts to maintain the virtual earth, D2 is reverse biased and so the output is just a 100 k resistor connected to 0 V. When Vin goes negative, the output rises positively, D2 is turned on and D1 turned off. As the virtual earth is being maintained, the output voltage is the exact inverse of the input voltage. This is true for all negative inputs. Therefore, the output is composed of positive going half sinewaves. Precision half wave rectification has occurred. In fact the diode error is very small, being equal to

$$\frac{600 \text{ mV}}{\text{(surplus voltage gain)}}$$

Therefore as the input frequency increases, and the surplus voltage gain decreases, precision falls.

By adding together the original and the half wave rectified signals together in the right ratio, it is possible to fill in the half cycle gaps and thus to generate precise full wave rectification. The addition of one summing op-amp and three resistors is all that is needed as shown opposite.

VOLTAGE TO CURRENT CONVERTER

The virtual earth of an op-amp and the current source characteristic of a transistor can be combined to produce a precision linear voltage to current converter. Consider the 'SOURCE' circuit. A positive voltage is applied and the op-amp adjusts itself so that a 'virtual earth' condition is maintained. This means that a current i flows through the input resistor R, where i = Vin/R. Now this current has to go somewhere, and so it flows through the PNP transistor and comes out of the collector and into its load. Thus, the input voltage generates a current which is linearly proportional to it. There are, however, three sources of error that will affect this linearity. First the input offset voltage of the op-amp may become significant at low levels of Vin. Second, the input bias current may well rob a lot of the current when Vin is low. Third, the base current of the transistor must be subtracted from the final output current. Note that the current gain of the transistor will change with collector current variations, and
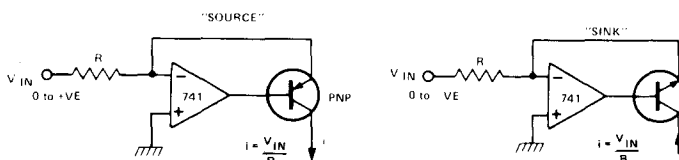

so the base current loss is not a fixed percentage. However, a precise voltage to current converter can be made using an op-amp with a FET input so that the bias current is low. Also, an input balance can be used to zero out the input offset voltage, and if a FET is used to replace the bipolar transistor, then the base current problem can be removed.

The 'SINK' circuit merely swaps the transistor to an npn type. Note that the input voltage now must be negative.

NON-INVERTING AMPLIFIER:

An op-amp is used to provide voltage gain, but in this case the output is in phase with the input. The minimum voltage is unity and occurs when RB is an open circuit. The op-amp has maximum bandwidth at unity gain, and any increase in the gain will cause a reciprocal decrease in bandwidth.

$$AV = + \frac{RA + RB}{RB}$$
$$Z_{IN} = RIN$$
$$Z_{OUT} = LOW$$
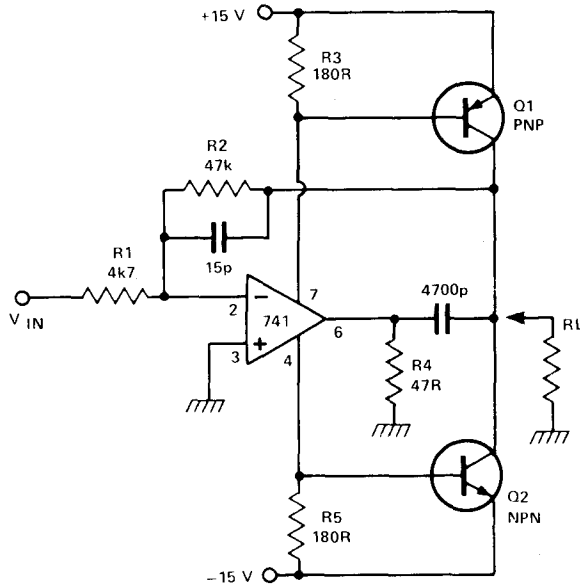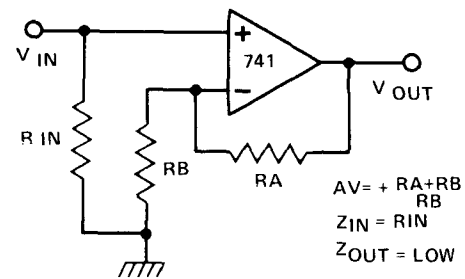
HIGH SLEW RATE AMPLIFIER

The slew rate of the op-amp has been increased by increasing the overall current generating capability, by the addition of a pair of transistors. These transistors increase the output voltage range by allowing the voltage to swing to within OV5 of either supply rails. The output of the op-amp hardly moves at all. Without an input signal, the output voltage is 0 V and the op-amp drains approximately 2 mA from the supply rails.

This current passes through the 180R resistors and sets up a voltage which is not quite sufficient to turn on either transistor. When a positive voltage is applied to the input, the op-amp tries to swing negative but it has a 47R (R4) resistor connected from its output to ground. Thus, as it tries to swing negative, it draws lots of current from the negative rail. This current flows through R5, and in doing so turns on Q2. This transistor then pulls R2 down and thus provides negative feedback. The same sequence of events occurs when the input is negative except that R3 and Q1 are then involved. Thus the high current capabilities of discrete transistors are combined with a high voltage gain of an op-amp to produce a moderately powerful amplifier. The voltage gain is set by R2/R1.

Transistors Q1 and Q2 introduce a phase shift, which may give rise to a high frequency instability and oscillation. This can be cured by some frequency compensation applied to the amplifier or by increasing the overall voltage gain.

SCHMITT TRIGGER

When dc positive feedback is applied around an op-amp, its output will come to rest in one of two states, that is in its most positive or most negative position. This type of circuit is known as a Schmitt Trigger and it is said to exhibit the property of hysterisis.

Consider the circuit shown in Figure 15. Let us assume that RB is 2 k and RA is 1 k and the output voltage is +10 V. Therefore the voltage at the non-inverting terminal is 3V3. When the input voltage becomes more positive than 3V3, the output of the op-amp will start to swing negative and in doing so will increase the voltage difference between the inputs. This will in turn make the output swing even more negative. Thus the process becomes regenerative, the output finally 'snapping' into its negative state (−10 V say). The only thing that will now change the op-amp's output is if the inverting input goes more negative than the non-inverting input. When this occurs it will revert back to its original state. The two input voltages at which these transitions happen are known as the upper and lower hysterisis levels.
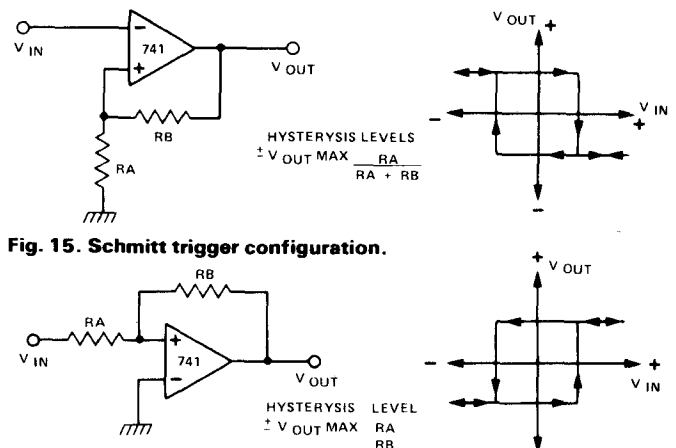
Fig. 15. Schmitt trigger configuration.

HYSTERYSIS LEVELS
$$\pm V_{OUT} MAX \frac{RA}{RA + RB}$$

HYSTERYSIS LEVEL
$$\pm V_{OUT} MAX \frac{RA}{RB}$$

HIGH PERFORMANCE SAMPLE AND HOLD

It is often necessary to have a circuit that will sample an analogue voltage and then remember it for a long time without any significant degradation of that voltage. This is known as a sample and hold circuit and one of its uses is to store the voltage from the keyboard connected to an electronic music synthesiser. The voltage is then used to control the pitch of a voltage controlled oscillator and so it is very important to have a high performance sample and hold. A drift of less than one semitone, (80 mV), in ten minutes is required. A sample and hold is simply an electronic switch, a storage capacitor and a high input impedance voltage follower. In the circuit shown, when switch SW1 is positive the FET is turned on, and has a resistance of about 400R. Thus the input voltage charges up the capacitor through the FET. When SW1 is negative, the FET is turned off, (pinched off), and can have a resistance of thousands of megohms. To get a long storage time the op-amp must have a very low input bias current. For the CA3140, this current is about 10 pico amps, i.e., 10^{-11} amps. Therefore the rate at which the capacitor will be discharged by this current can be worked out from the equation, $C(dv/dt) = i$
where dv/dt is the rate of change of voltage on the capacitor.

Therefore:

$$\frac{dv}{dt} = \frac{i}{C} = \frac{10^{-11}}{0.47 \times 10^{-6}} \quad 22\, uV/s$$

This is a very low drift rate, much better than we need. However, the actual drift rate will probably be in excess of this, due to surface leakage on the printed circuit board, leakage through the FET, and internal leakage in the capacitor. It is advisable to use a high voltage, non-polarised capacitor in this
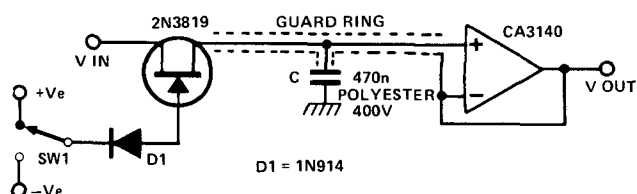

PRINTED CIRCUIT BOARD LAYOUT
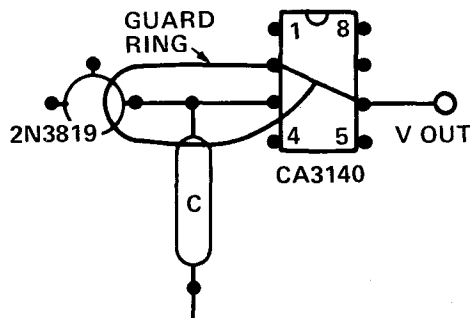
circuit to keep the leakage currents to a minimum. Also, to stop surface leakage a simple PCB trick can be used, that of making a guard ring around the sensitive components.

Normally any potential stored on the capacitor may leak to ground across the surface of the PCB, but if we make the surrounding surface a conducting track held at the same potential as that of the capacitor then the potential difference is virtually always zero, and hence the surface leakage is greatly reduced.

CLEANING UP DIGITALLY GENERATED SIGNALS WITH TWO SAMPLE AND HOLDS AND AN INTEGRATOR

The output from a digital to analogue converter (dac) is composed of a series of steps which have been selected by a series of binary numbers. The output of the dac may represent the result of some computation done by a microprocessor or the contents of a digital memory. If the number of bits that control the dac is low (less than eight), then the output will look like a series of discrete steps, plus lots of digital 'glitches'. Therefore, if this signal is to be displayed on an oscilloscope, the overall picture quality will be very poor. One way to clean up would be to join up all the steps with straight lines and if done successfully a great improvement can be obtained. The only problem is that the distance between steps is continuously varying and so the slope of the straight lines will need to be variable as well. This process is known as linear point interpolation and can be achieved with two sample and holds and an integrator.

A delayed gate pulse is generated, so that once the dac's output has settled, the sample and hold switches momentarily open, sample the information and then close. The output of the first sample and hold (IC1) drives an integrator (IC2), the output of which drives the second sample and hold (IC3). The second unit provides negative feedback around the integrator, but it is delayed by one time interval. Thus a momentary positive going signal will pass through the first sample and hold and cause the integrator to ramp in a negative direction. When the next time interval arrives, the first sample and hold returns to OV, and the second obtains a negative voltage. This then causes the integrator to ramp positively. The size of the
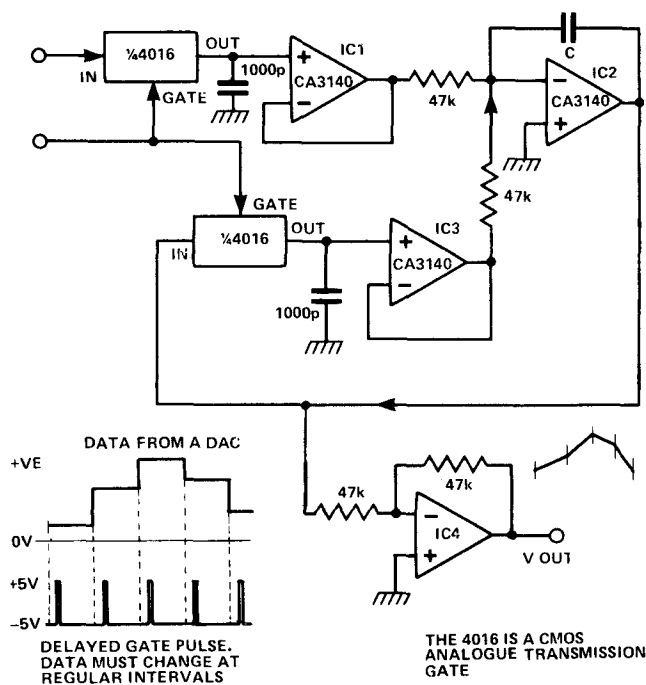
integrator's capacitor C should be chosen to suit the clock speed of the dac. An inverter, IC4 has been included to correct the invertion caused by the integrator.

FAST SYMMETRICAL ZENER CLAMPING

There are several problems with using zeners, back to back in series to get symmetrical clamping, the knee of the zener characteristics is rather sloppy, charge storage in the zeners causes speed problems and the zeners will have slightly different knee voltages so the symmetry will not be all that good. This circuit overcomes these problems. By putting the zener inside a diode bridge the same zener voltage is always experienced. The voltage errors due to the diodes are much smaller than those due to the zener. Also the charge storage of the bridge is much less. Lastly by biasing the zener on all the time, the knee appears to be much sharper.
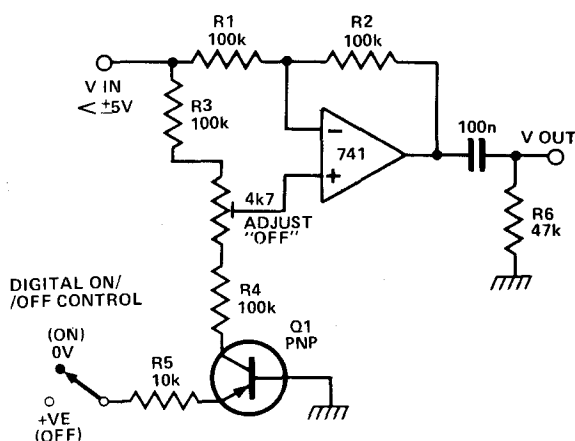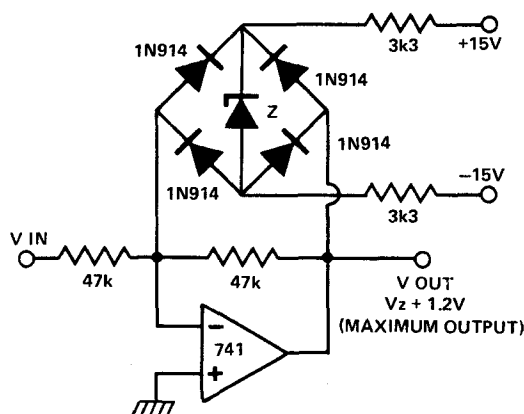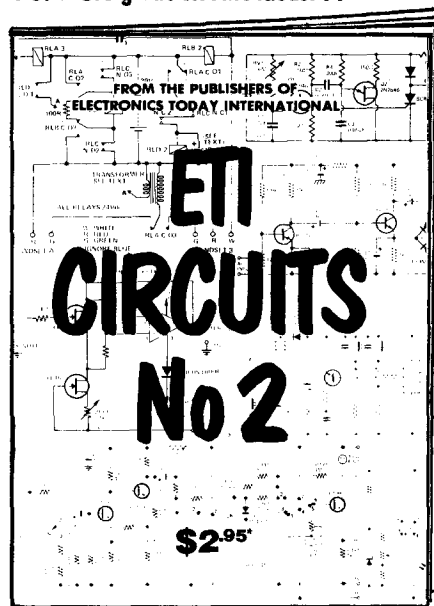
TRANSISTOR USED TO TURN AN OP AMP ON OR OFF

When transistor Q1 is switched off, the circuit behaves as a voltage follower. By applying a positive voltage to the emitter of Q1 via a 10 k resistor, the transistor is made to turn on and go into saturation. Thus the lower end of R4 is shorted to ground. The circuit has now changed into that of a differential amplifier (see fig. 7), but where the voltage difference is always 0 V. Now as long as the resistors in the two branches around the op amp are in the same ratio then there should be zero output. A 4k7 preset is used to null out any ratio errors so that the 'OFF' attenuation is more than 60 dB. The high common mode rejection ratio of a 741 enables this large attenuation to be obtained.

No Noise is Good Noise

The last op-amp characteristic to be discussed is 'Noise'. The noise figures given in the specifications are very confusing. This is due to the fact that noise is specified in so many different ways that it is often difficult to compare devices. One may be specified in terms of Equivalent Input Noise and another device in terms of $nV\sqrt{Hz}$ (nano volts per root Hertz)! As a generalisation it is true to say that most op-amps are relatively noisy. Some op-amps are labelled low noise, and these are quieter than the average op-amp but more noisy than a well designed discrete component amplifier. For audio work you can use ordinary op-amps for processing high level signals (100 mV to 3 V), but for amplifying low level signals (1 mV to 100 mV) you would be advised to use a low noise device. The larger the voltage gain you obtain from an op-amp stage, the worse will be the noise, therefore keep the closed loop gain to a bare minimum.

That is the end of the theory, now for some practical examples of op-amps in use.

THE 741 COOKBOOK

OPERATIONAL AMPLIFIERS (OP-AMPS) CAN be simply described as high-gain direct-coupled voltage amplifier 'blocks' that have a single output terminal but have both inverting and non-inverting input terminals. Op-amps can readily be used as inverting, non-inverting, and differential amplifiers in both a.c. and d.c. applications, and can easily be made to act as oscillators, tone filters, and level switches, etc.

Op-amps are readily available in integrated circuit form, and as such act as one of the most versatile building blocks available in electronics today. One of the most popular op-amps presently available is the device that is universally known as the "741" op-amp In this article we shall describe the basic features of this device, and show a wide variety of practical circuits in which it can be used.

BASIC OP-AMP CHARACTERISTICS AND CIRCUITS

In its simplest form, an op-amp consists of a differential amplifier followed by offset compensation and output stages, as shown in Fig. 1a. The differential amplifier has inverting and non-inverting input terminals, a high-impedance (constant current) tail to
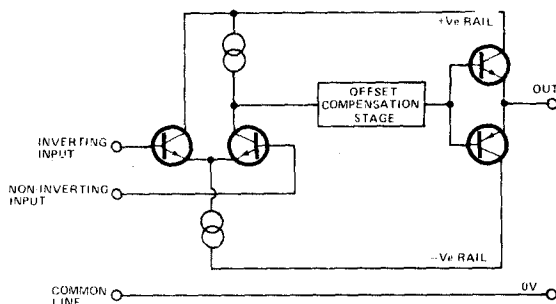

Fig. 1a Simplified op-amp equivalent circuit.

give a high input impedance and a high degree of common mode signal rejection. It also has a high-impedance (constant current) load to give a high degree of signal voltage stage gain.

The output of the differential amplifier is fed to a direct-coupled offset compensation stage, which

effectively reduces the output offset voltage of the differential amplifier to zero volts under quiescent conditions, and the output of the compensation stage is fed to a simple complementary emitter follower output stage, which gives a low output impedance.
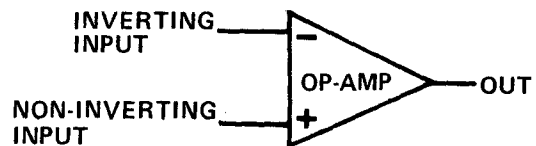

Fig.1b Basic op-amp symbol.

LINES OF SUPPLY

Op-amps are normally powered from split power supplies, providing +ve, −ve, and common (zero volt) supply rails, so that the output of the op-amp can swing either side of the zero volts value, and can be set at a true zero volts (when zero differential voltage is applied to the circuits input terminals.)

The input terminals can be used independently (with the unused terminal grounded) or simultaneously, enabling the device to function as an inverting, non-inverting, or differential amplifier. Since the device is direct-coupled throughout, it can be used to amplify both a.c. and d.c. input signals. Typically, they give basic low-frequency voltage gains of about 100 000 between input and output, and have input impedances of 1M or greater at each input terminal.

Fig. 1b shows the symbol that is commonly used to represent an op-amp, and 1c shows the basic supply connections that are used with the device. Note that both input and output signals of the op-amp are referenced to the ground or zero volt line.

SIGNAL BOX

The output signal voltage of the op-amp is proportional to the DIFFERENTIAL signal between its two input terminals, and is given by

$$e_{out} = A_0(e_1 - e_2)$$

where A_0 = the open-loop voltage gain of the op-amp (typically 100 000).

e_1 = signal voltage at the non-inverting input terminal.

e_2 = signal voltage at the inverting input terminal.

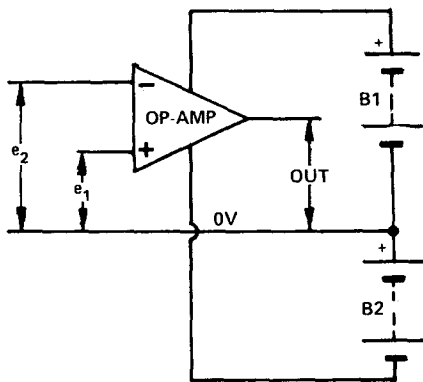Thus, if identical signals are simultaneously applied

Fig. 1c Basic supply connections of an op-amp.

Fig. 2b Transfer characteristics of the differential voltage comparator circuit.

to both input terminals, the circuit will (ideally) give zero signal output If a signal is applied to the inverting terminal only, the circuit gives an amplified and inverted output· If a signal is applied to the non-inverting terminal only, the circuit gives an amplified but non-inverted output.

By using external negative feedback components, the stage gain of the op-amp circuit can be very precisely controlled.
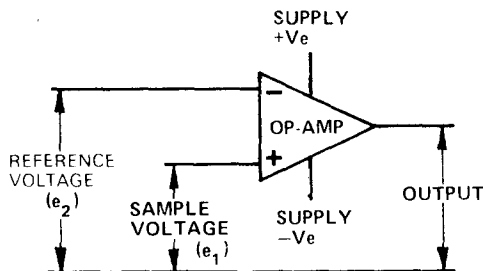
GOING TO GROUND

The op-amp can be made to function as a low-level inverting d.c. amplifier by simply grounding the non-inverting terminal and feeding the input signal to

Fig. 2a Simple differential voltage comparator circuit.
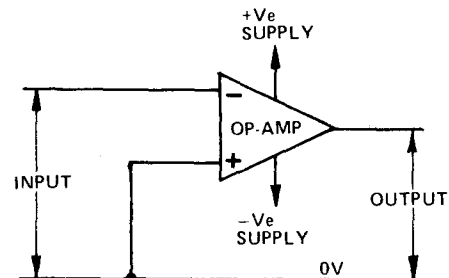
Fig. 3a Simple open-loop inverting d.c. amplifier.

TRANSFER REQUEST

Fig. 2a shows a very simple application of the op-amp. This particular circuit is known as a differential voltage comparator, and has a fixed reference voltage applied to the inverting input terminal, and a variable test or sample voltage applied to the non-inverting terminal. When the sample voltage is more than a few hundred microvolts above the reference voltage the op-amp output is driven to saturation in a positive direction, and when the sample is more than a few hundred microvolts below the reference voltage the output is driven to saturation in the negative direction.

Fig. 2b shows the voltage transfer characteristics of the above circuit. Note that it is the magnitude of the differential input voltage that dictates the magnitude of the output voltage, and that the absolute values of input voltage are of little importance. Thus, if a 1V reference is used and a differential voltage of only 200uV is needed to switch the output from a negative to a positive saturation level, this change can be caused by a shift of only 0.02% on a 1V signal applied to the sample input. The circuit thus functions as a precision voltage comparator or balance detector.

the inverting terminal, as shown in Fig. 3a. The op-amp is used 'open-loop' (without feedback) in this configuration, and thus gives a voltage gain of about 100 000 and has an input impedance of about 1M. The disadvantage of this circuit is that its parameters are dictated by the actual op-amp, and are subject to considerable variation between individual devices.

CLOSING LOOPS

A far more useful way of employing the op-amp is to use it in the closed-loop mode, i.e., with negative feedback. Fig. 3b shows the method of applying negative feedback to make a fixed-gain inverting d.c. amplifier. Here, the parameters of the circuit are controlled by feedback resistors R_1 and R_2. The gain, A of the circuit is dictated by the ratios of R_1 and R_2, and equals R_2/R_1.

The gain is virtually independent of the op-amp characteristics, provided that the open-loop gain (A_0) is large relative to the closed-loop gain (A). The input impedance of the circuit is equal to R_1, and again is virtually independent of the op-amp characteristics.

741 COOKBOOK

It should be noted at this point that although R_1 and R_2 control the gain of the complete circuit, they have no effect on the parameters of the actual op-amp, and the full open-loop gain of the op-amp is still available between its inverting input terminal and the output. Similarly, the inverting terminal continues to have a very high input impedance, and negligible signal current flows into the inverting terminal. Consquently, virtually all of the R_1 signal current also flows in R_2, and
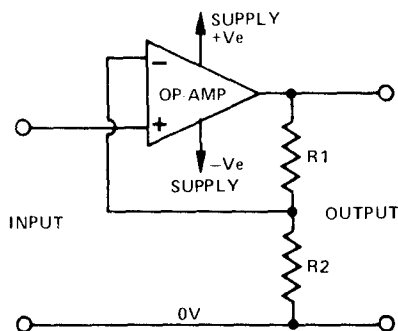

Fig. 4a Basic non-inverting d.c. amplifier

signal currents i_1 and i_2 can be regarded as being equal, as indicated in the diagram.

Since the signal voltage appearing at the output terminal end of R_2 is A times greater than that appearing at the inverting terminal end, the current flowing in R_2 is A times greater than that caused by the inverting terminal signal only. Consequently, R_2 has an apparent value of R_2/A when looked at from its inverting terminal end, and the R_1-R_2 junction thus appears as a low-impedance VIRTUAL GROUND point.
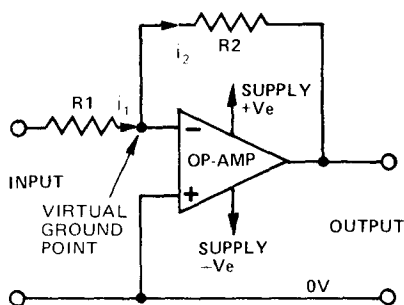

Fig. 3b Basic closed-loop inverting d.c. amplifier.

INVERT OR NOT TO INVERT...

It can be seen from the above description that the Fig. 3b circuit is very versatile. Its gain and input impedance can be very precisely controlled by suitable choice of R_1 and R_2, and are unaffected by variations in the op-amp characteristics. A similar thing is true of the non-inverting d.c. amplifier circuit shown in Fig. 4a. In this case the voltage gain is equal to $(R_1 + R_2)/R_2$ and the input impedance is approximately equal to $(A_0/A)Zin$ where Zin is the open-loop input impedance of the op-amp. A great advantage of this circuit is that it has a very high input impedance.

FOLLOW THAT VOLTAGE

The op-amp can be made to function as a precision voltage follower by connecting it as a unity-gain non-inverting d.c. amplifier, as shown in Fig. 4b. In this case the input and output voltages of the circuit are identical, but the input impedance is very high and is roughly equal to $A_0 \times Z_{in}$.

The basic op-amp circuits of Figs. 2a to 4b are shown as d.c. amplifiers, but can readily be adapted for a.c. use. Op-amps also have many applications other than as simple amplifiers. They can easily be made to function as precision phase splitters, as adders or subtractors, as active filters or selective amplifiers, as precision half-wave or full-wave rectifiers, and as oscillators or multivibrators, etc.
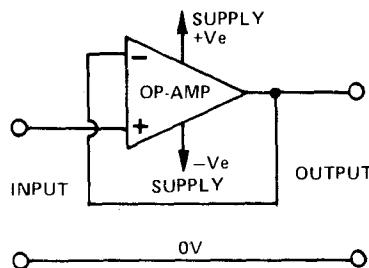

Fig. 4b Basic unity-gain d.c. voltage follower

OP-AMP PARAMETERS

An ideal op-amp would have an infinite input impedance, zero output impedance, infinite gain and infinite bandwidth, and would give perfect tracking between input and output. Practical op-amps fall far short of this ideal, and have finite gain, bandwidth, etc., and give tracking errors between the input and output signals. Consequently, various performance parameters are detailed on op-amp data sheets, and indicate the measure of "goodness" of the particular device. The most important of these parameters are detailed below.

OPEN-LOOP VOLTAGE GAIN, A_0. This is the low-frequency voltage gain occuring directly between the input and output terminals of the op-amp, and may be expressed in direct terms or in terms of dB. Typically, d.c. gain figures of modern op-amps are 100 000, or 100dB.

INPUT IMPEDANCE, Z_{in}. This is the impedance looking directly into the input terminals of the op-amp when it is used open-loop, and is usually expressed in terms of resistance only. Values of 1M are typical of modern op-amps with bi-polar input stages, while F.E.T. input types have impedances of a million meg or greater.

OUTPUT IMPEDANCE, Z_0. This is the output impedance of the basic op-amp when it is used open-loop, and is usually expressed in terms of resistance only. Values of a few hundred ohms are typical of modern op-amps.

INPUT BIAS CURRENT, I_b. Many op-amps use bipolar transistor input stages, and draw a small bias current from the input terminals. The magnitude of this current is denoted by I_b, and is typically only a fraction of a microamp.

741 COOKBOOK

SUPPLY VOLTAGE RANGE, V$_s$. Op-amps are usually operated from two sets of supply rails, and these supplies must be within maximum and minimum limits. If the supply voltages are too high the op-amp may be damaged, and if the supply voltages are too low the op-amp will not function correctly. Typical supply limits are $\pm 3V$ to $\pm 15V$.

INPUT VOLTAGE RANGE, V$_{i(max)}$. The input voltage to the op-amp must never be allowed to exceed the supply line voltages, or the op-amp may be damaged. V$_{i(max)}$ is usually specified as being one or two volts less than v$_s$.

OUTPUT VOLTAGE RANGE, V$_{o(max)}$. If the op-amp is over driven its output will saturate and be limited by the available supply voltages, so V$_{o(max)}$ is usually specified as being one or two volts less than V$_s$.

DIFFERENTIAL INPUT OFFSET VOLTAGE, V$_{io}$. In the ideal op-amp perfect tracking would exist between the input and output terminals of the device, and the output would register zero when both inputs were grounded. Actual op-amps are not perfect devices, however, and in practice slight imbalances exist within their input circuitry and effectively cause a small offset or bias potential to be applied to the input terminals of the op-amp. Typically, this DIFFERENTIAL INPUT OFFSET VOLTAGE has a value of only a few millivolts, but when this voltage is amplified by the gain of the circuit in which the op-amp is used it may be sufficient to drive the op-amp output to saturation. Because of this, most op-amps have some facility for externally nulling out the offset voltage.

COMMON MODE REJECTION RATION, c.m.r.r. The ideal op-amp produces an output that is proportional to the difference between the two signals applied to its input terminals, and produces zero output when identical signals are applied to both inputs simultaneously, i.e., in common mode. In practical op-amps, common mode signals do not entirely cancel out, and produce a small signal at the op-amps output terminal. The ability of the op-amp to reject common mode signals is usually expressed in terms of common mode rejection ratio, which is the ratio of the op-amps gain with differential signals to the op-amps gain with common mode signals. C.m.r.r. values of 90dB are typical of modern op-amps.

TRANSITION FREQUENCY, f$_T$. An op-amp typically gives a low-frequency voltage gain of about 100dB, and in the interest of stability its open-loop frequency response is tailored so that the gain falls off as the frequency rises, and falls to unity at a transition frequency denoted f$_T$. Usually, the response falls off at a rate of 6dB per octave or 20dB per decade. Fig. 5 shows the typical response curve of the type 741 op-amp, which has an f$_T$ of 1MHz and a low frequency gain of 100dB.

Note that, when the op-amp is used in a closed-loop amplifier circuit, the bandwidth of the circuit depends on the closed-loop gain. If the amplifier is used to give a gain of 60dB its bandwidth is only 1kHz, and if it is used to give a gain of 20dB its bandwidth is 100kHz.

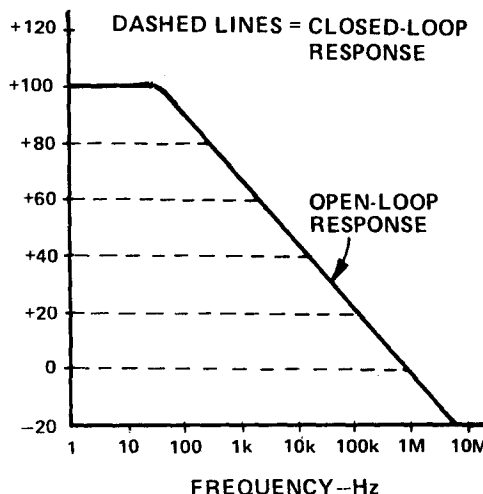The f$_T$ figure can thus be used to represent a gain-bandwidth product.

Fig. 5 Typical frequency response curve of the 741 op-amp.

	PARAMETER	741 VALUE
A$_o$	OPEN-LOOP VOLTAGE GAIN	100dB
Z$_{IN}$	INPUT IMPEDANCE	1M
Z$_o$	OUTPUT IMPEDANCE	150R
I$_b$	INPUT BIAS CURRENT	200nA
V$_s$ (MAX)	MAXIMUM SUPPLY VOLTAGE	+18V
V$_i$ (MAX)	MAXIMUM INPUT VOLTAGE	+13V
V$_o$ (MAX)	MAXIMUM OUTPUT VOLTAGE	+14V
V$_{io}$	DIFFERENTIAL INPUT OFFSET VOLTAGE	2mV
c.m.m.r.	COMMON MODE REJECTION RATIO	90dB
F$_T$	TRANSITION FREQUENCY	1MHZ
S	SLEW RATE	1V/uS

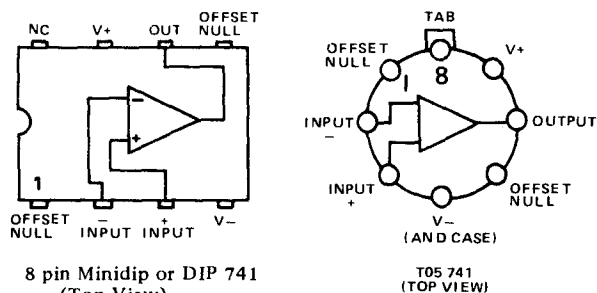*Table 1 Typical characteristics of the 741 op-amp.*

SLEW RATE. As well as being subject to normal bandwidth limitations, op-amps are also subject to a phenomenon known as slew rate limiting, which has the effect of limiting the maximum rate of change of voltage at the output of the device. Slew rate is normally specified in terms of volts per microsecond, and values in the range 1V/us to 10V/us are common with most popular types of op-amp. One effect of slew rate limiting is to make a greater bandwidth available to small output signals than is available to large output signals.

THE 741 OP-AMP.

Early types of i.c. op-amp, such as the well known 709 type, suffered from a number of design weaknesses. In particular, they were prone to a phenomenon known as INPUT LATCH-UP, in which the input circuitry tended to switch into a locked state if special precautions were not taken when connecting the input signals to the input terminals, and tended to self-destruct if a short circuit were inadvertently placed across the op-amp output terminals. In addition, the op-amps were prone to bursting into unwanted oscillations when used in the linear amplifier mode, and required the use of external frequency compensation components for stability control.

These weaknesses have been eliminated in the type 741 op-amp. This device is immune to input latch-up problems, has built-in output short circuit protection, and does not require the use of external frequency compensation components. The typical performance

741 COOKBOOK

Fig. 6 Outlines and pin connections of the two most popular 741 packages.

characteristics of the device are listed in Table 1.

The type 741 op-amp is marketed by most i.c. manufacturers, and is very readily available. Fig. 6 shows the two most commonly used forms of packaging of the device. Throughout this chapter, all practical circuits are based on the standard 8-pin dual-in-line (D.I.L. or DIP) version of the 741 op-amp.
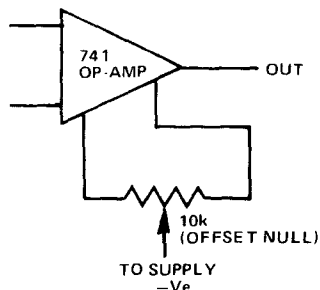

Fig. 7 Method of applying offset nulling to the 741 op-amp.

The 741 op-amp can be provided with external offset nulling by wiring a 10k pot between its two null terminals and taking the pot slider to the negative supply rail, as shown in Fig. 7.

Having cleared up these basic points, let's now go on and look at a range of practical applications of the 741 op-amp.
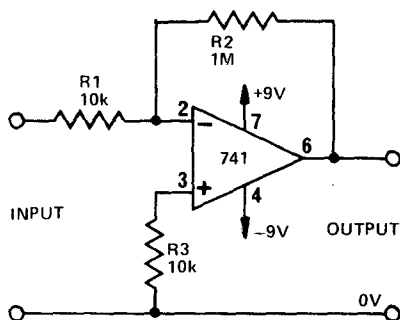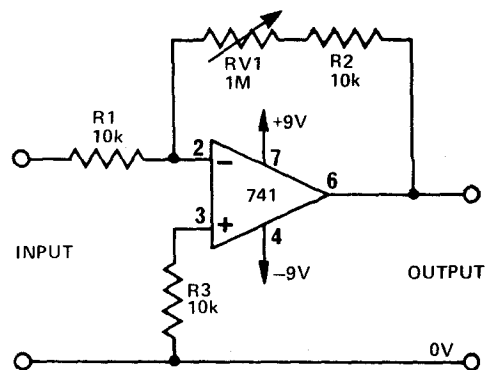

Fig. 8a x100 inverting d.c. amplifier.

Fig. 8b Variable gain (x1 to x100) inverting d.c. amplifier.

BASIC LINEAR AMPLIFIER PROJECTS. (Figs. 8 to 11).

Figs. 8 to 11 show a variety of ways of using the 741 in basic linear amplifier applications.

The 741 can be made to function as an inverting amplifier by grounding the non-inverting input terminal and feeding the input signal to the inverting terminal. The voltage gain of the circuit can be precisely controlled by selecting suitable values of external feedback resistance. Fig. 8a shows the practical connections of an inverting d.c. amplifier with a pre-set gain of x100. The voltage gain is determined by the ratios of R_1 and R_2, as shown in the diagram.

The gain can be readily altered by using alternative R_1 and/or R_2 values. If required, the gain can be made variable by using a series combination of a fixed and a variable resistor in place of R_2, as shown in the circuit of Fig. 8b, in which the gain can be varied over the range x1 to x100 via R_2.

VARIATIONS

A variation of the basic inverting d.c. amplifier is shown in Fig. 9a. Here, the feedback connection to R_2 is taken from the output of the R_3–R_4 output potential divider, rather than directly from the output of the op-amp, and the voltage gain is determined by the ratios of this divider as well as by the values of R_1 and
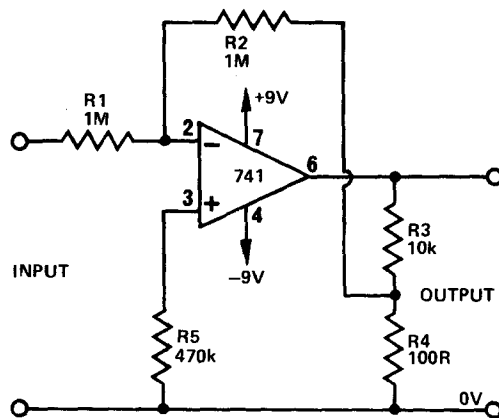

Fig. 9a High impedance x100 inverting d.c. amplifier.

R_2. The important feature of this circuit is that it enables R_1, which determines the input impedance of the circuit, to be given a high value if required, while at the same time enabling high voltage gain to be achieved.

The basic inverting d.c. amplifier can be adapted for

741 COOKBOOK

a.c. use by simply wiring blocking capacitors in series with its input and output terminals, as shown in the x100 inverting a.c. amplifier circuit of Fig. 9b.
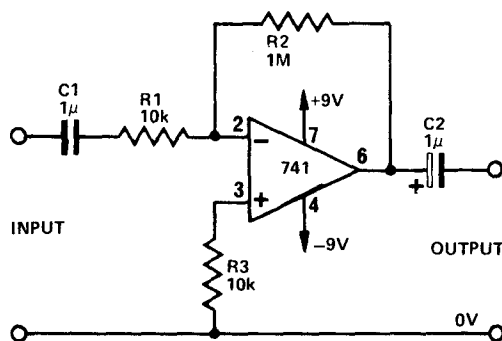

Fig. 9b x100 inverting a.c. amplifier.

NON-INVERTING . . .

The amp can be made to function as a non-inverting amplifier by feeding the input signal to its non-inverting terminal and applying negative feedback to the inverting terminal via a resistive potential divider that is connected across the op-amp output. Fig. 10a shows the connections for making a fixed gain (x100) d.c. amplifier.

The voltage gain of the Fig. 10a circuit is determined by the ratios of R_1 and R_2. If R_2 is given a value of zero the gain falls to unity, and if R_1 is given a value of zero the gain rises towards infinity (but in practice is limited to the open-loop gain of the op-amp). If required, the gain can be made variable by replacing R_2 with a
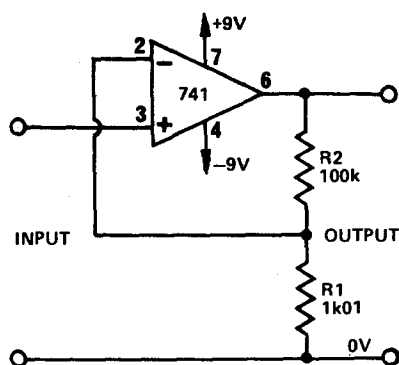

Fig. 10a Non-inverting x100 d.c. amplifier.

potentiometer and connecting the pot slider to the inverting terminal of the op-amp, as shown in the circuit of Fig. 10b. The gain of this circuit can be varied over the range x1 to x100 via R_1.

. . . AND RESISTANCE TO INPUTS

A major advantage of the non-inverting d.c. amplifier is that it has a very high input resistance. In theory, the input resistance is equal to the open-loop input resistance (typically 1M) multiplied by the open-loop voltage gain (typically 100 000) divided by the actual circuit voltage gain. In practice, input resistance values of hundreds of megohms can readily be obtained.
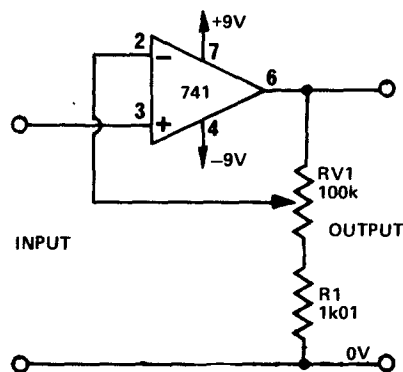

Fig. 10b Non-inverting variable gain (x1 to x100) d.c. amplifier.

BLOCKING OUT

The basic non-inverting d.c. circuit of Fig. 10 can be modified to operate as a.c. amplifiers in a variety of ways. The most obvious approach here is to simply wire blocking capacitors in series with the inputs and outputs, but in such cases the input terminal must be d.c. grounded via a suitable resistor, as shown by R_3 in the non-inverting x100 a.c. amplifier of Fig. 11a. If this resistor is not used the op-amp will have no d.c. stability, and its output will rapidly drift into saturation. Clearly, the input resistance of the Fig. 11a circuit is equal to R_3, and R_3 must have a relatively low value in the interest of d.c. stability. This circuit thus loses the non-inverting amplifier's basic advantage of high input resistance.
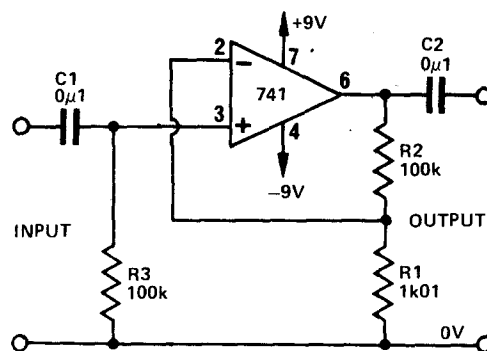
Fig. 11a Non-inverting, high input-impedance, x100 a.c. amplifier.

DRIFTING INTO STABILITY

A useful development of the Fig. 11a circuit is shown in Fig. 11b. Here, the values of R_1 and R_2 are increased and a blocking capacitor is interposed between them. At practical operating frequencies this capacitor has a negligible impedance, so the voltage gain is still determined by the ratios of the two resistors. Because of the inclusion of the blocking capacitor, however, the inverting terminal of the op-amp is subjected to virtually 100% d.c. negative feedback from the output terminal of the op-amp, and the circuit thus has excellent d.c. stability. The low end of R_3 is

741 COOKBOOK

connected to the C_3–R_1 junction, rather than directly to the *ground line*, and the signal voltage appearing at this point is virtually identical with that appearing at the non-inverting terminal of the op-amp·

Fig. 11b   Non-inverting x100 a.c. amplifier.

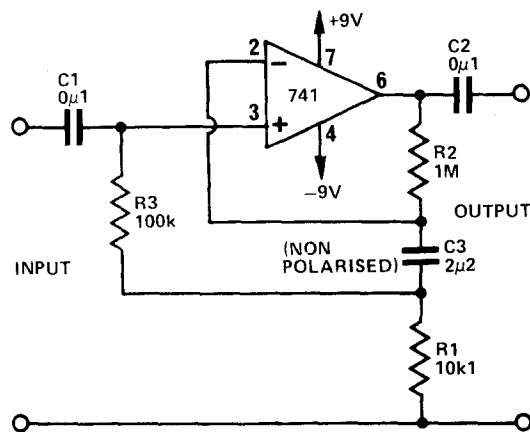Consequently, identical signal voltages appear at both ends of R_3, and the apparent impedance of this resistor is increased close to infinity by bootstrap action.

This circuit thus has good d.c. stability and a very high input impedance· In practice, this circuit gives a typical input impedance of about 50M.

VOLTAGE FOLLOWER PROJECTS (Figs. 12 to 13).

A 741 can be made to function as a precision voltage follower by connecting it as a unity-gain non-inverting amplifier. Fig. 12a shows the practical connections for making a d.c. voltage follower. Here, the input signal is applied directly to the non-inverting terminal of the op-amp, and the inverting terminal is connected directly to the output, so the circuit has 100% d.c. negative feedback and acts as a unity-gain non-inverting d.c. amplifier.

The output signal voltage of the circuit is virtually identical to that of the input, so the output is said to 'follow' the input voltage. The great advantage of this circuit is that it has a very high input impedance (as high as hundreds of megohms) and a very low output impedance (as low as a few ohms). The circuit acts effectively as an impedance transformer.
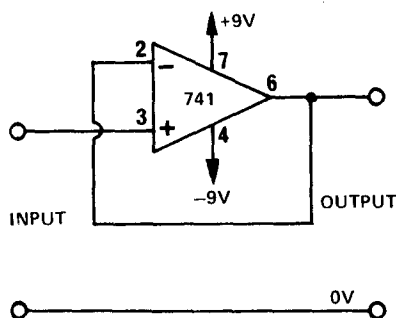

Fig. 12a d.c. voltage follower.

PRACTICE, AND ITS LIMITS

In practice the output of the basic Fig. 12a circuit will follow the input to within a couple of millivolts up to magnitudes within a volt or so of the supply line potentials. If required, the circuit can be made to follow to within a few microvolts by adding the offset null facility to the op-amp.

The d.c. voltage follower can be adapted for a.c. use by wiring blocking capacitors in series with its input and output terminals and by d.c.-coupling the non-inverting terminal of the op-amp to the zero volts line via a suitable resistor, as shown by R_1 in Fig. 12b. R_1 should have a value less than a couple of megohms, and restricts the available input impedance of the voltage follower.

LACED UP OHMS

If a very high input-impedance a.c. voltage follower is needed, the circuit of Fig. 12c can be used. Here, R_1 is boostrapped from the output of the op-amp, and its apparent impedance is greatly increased. This circuit has a typical impedance of hundreds of megohms.
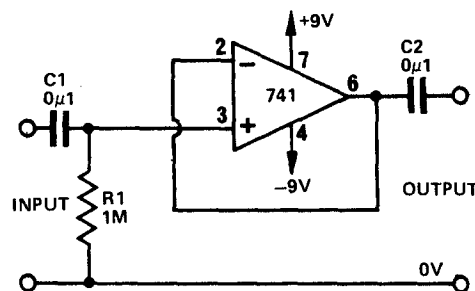
Fig. 12b a.c. voltage follower.
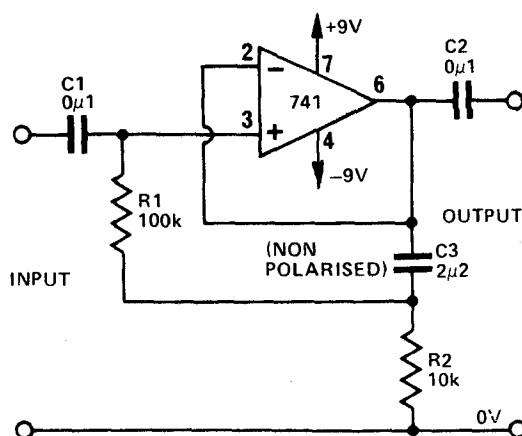
Fig. 12c Very high input-impedance a.c. voltage follower.

DRIVING CIRCUITS AMP-LY

The 741 op-amp is capable of providing output currents up to about 5mA, and this is consequently the current-driving limit of the three voltage follower circuits that we have looked at so far. The current-driving capabilities of the circuits can readily be increased by wiring simple or complementary emitter follower booster stages between the op-amp output terminals and the outputs of the actual circuits, as shown in Figs. 13a and 13b respectively.

Note in each case that the base-emitter junction(s) of

741 COOKBOOK

the output transistor(s) are included in the negative feedback loop of the circuit. Consequently, the 600mV knee voltage of each junction is effectively reduced by a factor equal to the open-loop gain of the op-amp, so the junctions do not adversely effect the voltage-following characteristics of either circuit.

The Fig. 13a circuit is able to source current only, and can be regarded as a unidirectional, positive-going, d.c. voltage follower. The Fig. 13b circuit can both source and sink output currents, and thus gives bidirectional follower action. Each circuit has a current-driving capacity of about 50mA. This figure is dictated by the limited power rating of the specified output transistors. The drive capability can be increased by using alternative transistors.
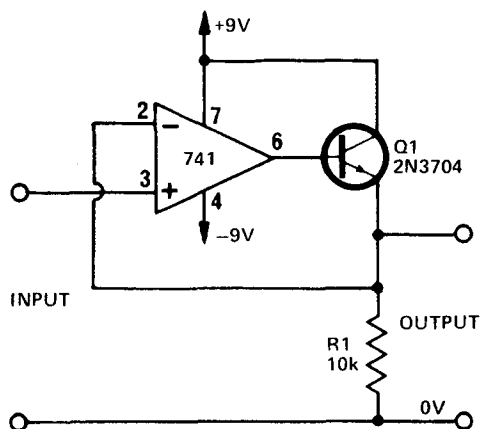

Fig. 13a Unidirectional d.c. voltage follower with boosted output (variable from 0V to +8V at 50mA.)
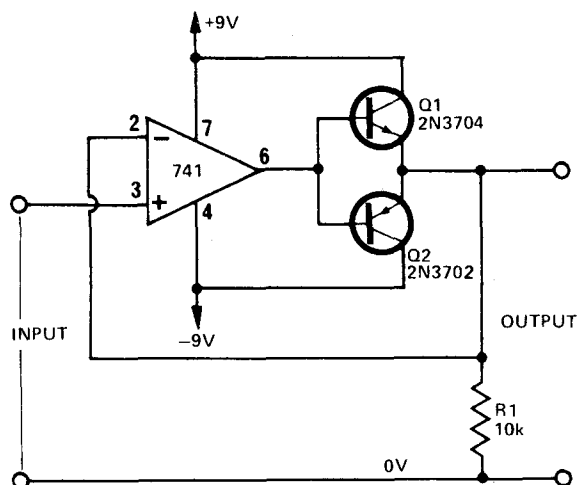

Fig. 13b Bidirectional d.c. voltage follower with boosted output (variable from 0V to± ±8V at 50mA).

MISC AMP PROJECTS (Figs. 14 to 22)

Figs. 14 to 22 show a miscellaneous assortment of 741 amplifier projects, ranging from d.c. adding circuits to frequency-selective amplifiers.

Fig. 14 shows the circuit of a unity-gain inverting d.c. adder, which gives an output voltage that is equal to the sum of the three input voltages. Here, input resistors R_1 to R_3 and feedback resistor R_4 each have the same value, and the circuit thus acts as a unity-gain inverting d.c. amplifier between each input terminal and the output. Since the current flowing in each input resistor also flows in feedback resistor R_4, the total current flowing in R_4 is equal to the sum of the input currents, and the output voltage is equal to the negative sum of the input voltages. The circuit is shown with only three input connections, but in fact can be provided with any number of input terminals. The circuit can be made to function as a so-called 'audio mixer' by wiring blocking capacitors in series with each input terminal and with the output terminal.
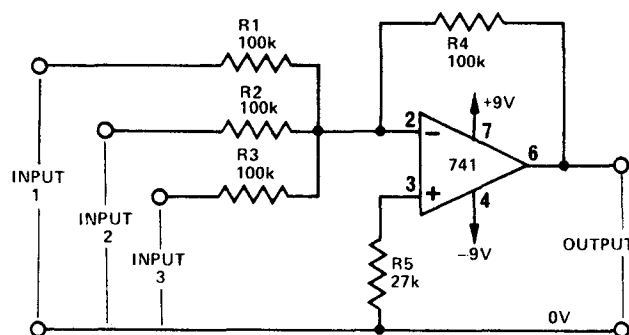
Fig. 14 Unity-gain inverting d.c. adder, or 'audio mixer'.

FIG. 15 shows how two unity-gain inverting d.c. amplifiers can be wired in series to make a precision unity-gain balanced phase-splitter. The output of the first amplifier is an inverted version of the input signal, and the output of the second amplifier is a non-inverted version.
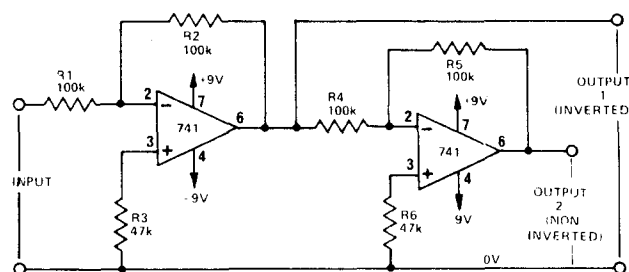
Fig. 15 Unity-gain balanced d.c. phase-splitter.

FIG. 16 shows how a 741 can be used as a unity-gain differential d.c. amplifier. The output of this circuit is equal to the difference between the two input signals or voltages, or to e_1-e_2. Thus, the circuit can also be used as a subtractor. In this type of circuit the component values are chosen such that $R_1/R_2 = R_4/R_3$, in which case the voltage gain $A_v = R_2/R_1$. The circuit can thus be made to give voltage gain if required.

741 COOKBOOK
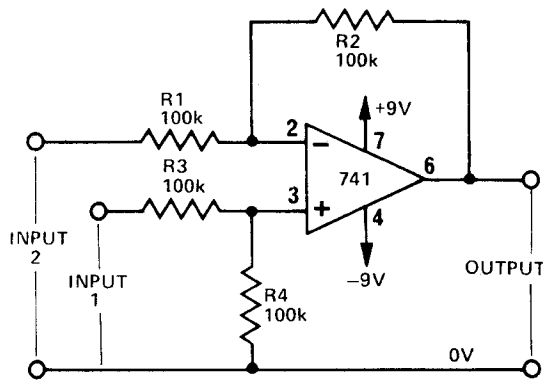
Fig. 16 Unity-gain differential d.c. amplifier, or subtractor.

FIG. 17 shows the amp can be made to act as a non-linear (semi-log) a.c. voltage amplifier by using a couple of ordinary silicon diodes as feedback elements. The voltage gain of the circuit depends on the magnitude of applied input signal, and is high when input signals are low, and low when input signals are high. The measured performance of the circuit is shown in the table, and can be varied by using alternative R_1 values.

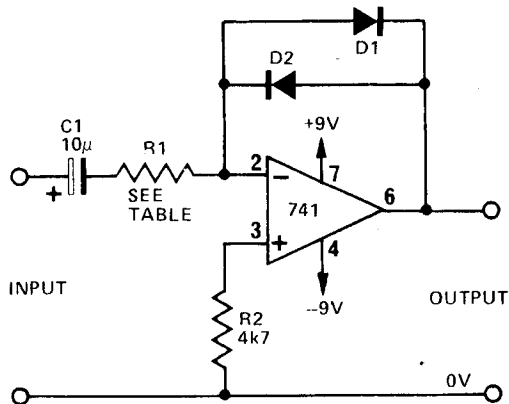| INPUT VOLTS (RMS)	$R_1 = 1k\Omega$		$R_1 = 10k\Omega$	
	V_{out}(R.M.S.)	GAIN	V_{out}(RMS)	GAIN
1 mV	110 mV	X 110	21 mV	X 21
10 mV	330 mV	X 33	170 mV	X 17
100 mV	450 mV	X4.5	360 mV	X 3.6
1V	560 mV	X0.56	470 mV	X 0.47
10V	600 mV	X0.07	560 mV	X0.056

Fig. 17 Circuit and performance table of non-linear (semi-log) a.c. voltage amplifier.

FIG. 18 shows how the 741 can be used together with a junction-type field-effect transistor (JFET) to make a so-called constant-volume amplifier. The action of this type of circuit is such that its peak output voltage is held sensibly constant, without distortion, over a wide range of input signal levels, and this particular circuit gives a sensibly constant output over a 30dB range of input signal levels.

The measured performance of the circuit is shown in the table. C_1 determines the response time of the
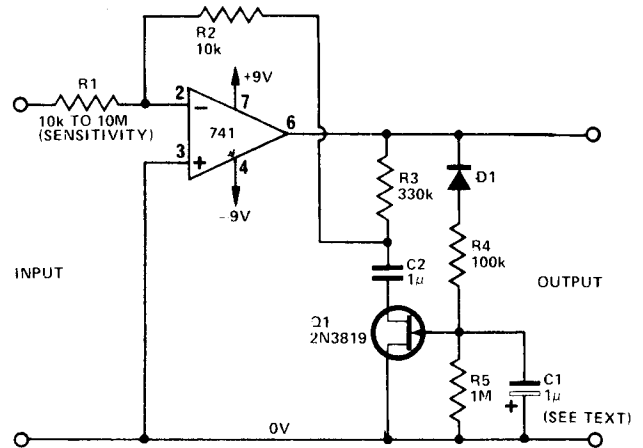
V_{IN} ($R_1 = 10K\Omega$)	V_{IN} ($R_1 = 100K\Omega$)	V_{IN} ($R_1 = 1M\Omega$)	V_{IN} ($R_1 = 10M\Omega$)	V_{OUT}
50 mV	500 mV	5 V	50 V	2.85 V
20 mV	200 mV	2 V	20 V	2.81 V
10 mV	100 mV	1 V	10 V	2.79 V
5 mV	50 mV	500 mV	5 V	2.60 V
2 mV	20 mV	200 mV	2 V	2.03 V
1 mV	10 mV	100 mV	1 V	1.48 V
500 μV	5 mV	50 mV	500 mV	0.89 V
200 μV	2 mV	20 mV	200 mV	0.40 V
100 μV	1 mV	10 mV	100 mV	0.20 V
50 μV	500 μV	5 mV	50 mV	0.10 V

Fig. 18 Circuit and performance details of constant-volume amplifier.

ACTION TAKEN

The action of the Fig. 18 circuit relies on the fact that the JFET can act as a voltage-controlled resistance which appears as a low value when zero bias is applied to its gate and as a high resistance when its gate is negatively biased. The JFET and R_3 act as a gain-determining a.c. voltage divider (via C_2), and the bias to the JFET gate is derived from the circuits output via the D_1-C_1 network. When the circuit output is low the JFET appears as a low resistance, and the op-amp gives high voltage gain.

When the circuit output is high the JFET appears as a high resistance, and the op-amp gives low voltage gain. The output level of the circuit is thus held sensibly constant by negative feedback.
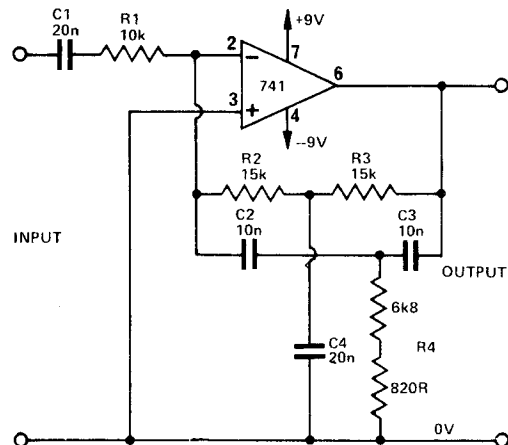

Fig. 19 1kHz tuned (acceptor) amplifier (twin-T).

741 COOKBOOK

CHOOSE YOUR FREQUENCY

The 741 op-amp can be made to function as a frequency-selective amplifier by connecting frequency-sensitive networks into its feedback loops. Fig. 19 shows how a twin-T network can be connected to the op-amp so that it acts as a tuned (acceptor) amplifier, and Fig. 20 shows how the same twin-T network can be connected so that the op-amp acts as a notch (rejector) filter. The values of the twin-T network are chosen such that $R_2 = R_3 = 2 \times R_4$, and $C_2 = C_4/2$, in which case its centre (tuned) frequency $= 1/6.28\ R_2.C_2$. With the component values shown, both circuits are tuned to approximately 1kHz.
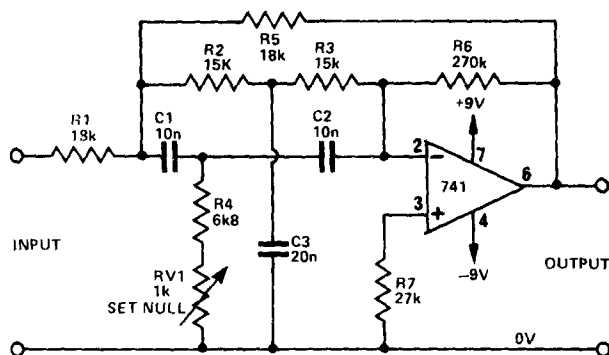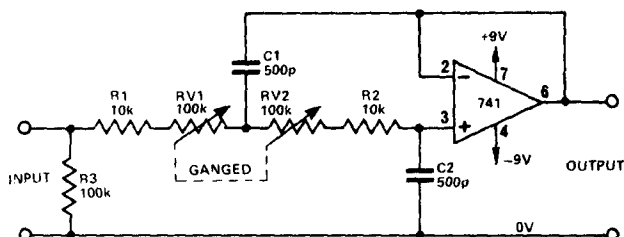

Fig. 20 1kHz notch (reject) filter.

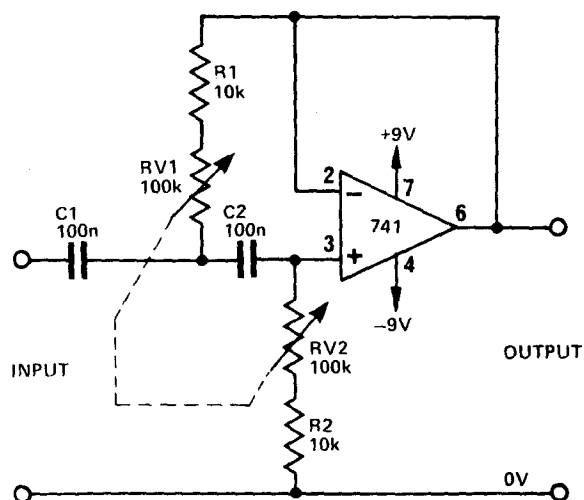Fig. 21 Variable low-pass filter, covering 2.2kHz to 24kHz.

Fig. 22 Variable high-pass filter, covering 235Hz to 2.8kHz.

Finally, to complete this section, Figs. 21 and 22 show the circuits of a couple of variable-frequency audio filters. The Fig. 21 circuit is that of a low-pass filter which covers the range 2.2kHz to 24kHz, and the Fig. 22 circuit is that of a high-pass filter which covers the range 235Hz to 2.8kHz. In each case, the circuit gives unity gain to signals beyond its cut-off frequency, and gives a 2nd order response (a change of 12dB per octave) to signals within its range.

INSTRUMENTATION PROJECTS (Figs. 23 to 31)

Figs. 23 to 31 show a variety of instrumentation projects in which the 741 can be used. The circuits range from a simple voltage regulator to a linear-scale ohmmeter.
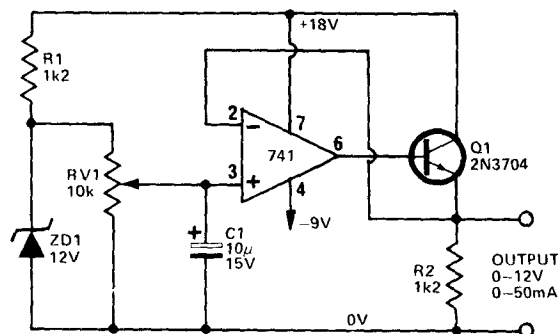

Fig. 23 Simple variable-voltage supply.

FIG. 23 shows the circuit of a simple variable-voltage power supply, which gives a stable output that is fully adjustable from 0V to 12V at currents up to a maximum of about 50mA. The operation of the circuit is quite simple. ZD_1 is a zener diode, and is energised from the positive supply line via R_1. A constant reference potential of 12V is developed across the zener diode, and is fed to variable potential divider RV_1.

The output of this divider is fully variable from 0V to 12V, and is fed ·to the non-inverting input of the op-amp. The op-amp is wired as a unity-gain voltage follower, with Q_1 connected as an emitter follower current-booster stage in series with its output.

Thus, the output voltage of the circuit follows the voltage set at the op-amp input via RV_1, and is fully variable from 0V to 12V. Note that the circuit uses an 18V positive supply and a 9V negative supply.

Also note that the voltage range of the above circuit can be increased by using higher zener and unregulated supply voltages, and that its current capacity can be increased by using one or more power transistors in place of Q_1.

FIG. 24 shows how a 741 op-amp can be used as the basis of a stabilised power supply unit (P.S.U.) that covers the range 3V to 30V at currents up to 1A. Here, the voltage supply to the op-amp is stabilized at 33V via ZD_1, and a highly temperature-stable reference of 3V is fed to the input of the op-amp via ZD_2.

The op-amp and output transistors Q_1-Q_2 are wired as a variable-gain non-inverting d.c. amplifier, with gain variable from unity to x10 via RV_1, and the output voltage is thus fully variable from 3V to 30V via RV_1. The output voltage is fully stabilized by negative feedback.

25

741 COOKBOOK

FIG. 26a shows how a 741 can be used in conjunction with a couple of silicon diodes as a precision half-wave rectifier. Conventional diodes act as imperfect rectifiers of low-level a.c. signals, because they do not begin to conduct significantly until the applied signal voltage exceeds a 'knee' value of about 600mV.

When diodes are wired into the negative feedback loop of the circuit as shown the 'knee' voltage is effectively reduced by a factor equal to the open-loop gain of the op-amp, and the circuit thus acts like a near-perfect rectifier.

The overall voltage gain of the Fig. 26a circuit is dictated by the ratios of R_1 and R_2 to R_3, as in the case of a conventional inverting amplifier, and this circuit thus gives a gain of unity. The circuit can be made to
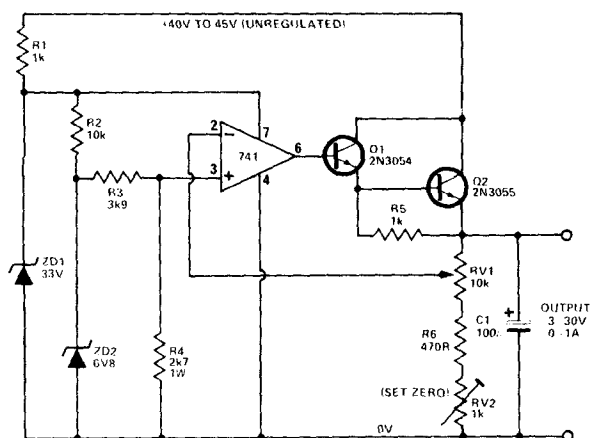
Fig. 24 3V - 30V, 0-1 amp stabilised p.s.u.

FIG. 25 shows how overload protection can be applied to the above circuit. Here, current-sensing resistor R_7 is wired in series with the output of the regulator, and cut-out transistor Q_3 is driven from this resistor and is wired so that its base-collector junction is able to short the base-emitter junction of the Q_1-Q_2 output transistor stage.

Normally, Q_3 is inoperative, and has no effect on the circuit, but when P.S.U. output currents exceed 1A a potential in excess of 600mV is developed across R_7 and biases Q_3 on, thus causing Q_3 to shunt the base-emitter junction of the Q_1-Q_2 output stage and hence reducing the output current. Heavy negative feedback takes place in this action, and the output current is automatically limited to 1A, even under short-circuit conditions.
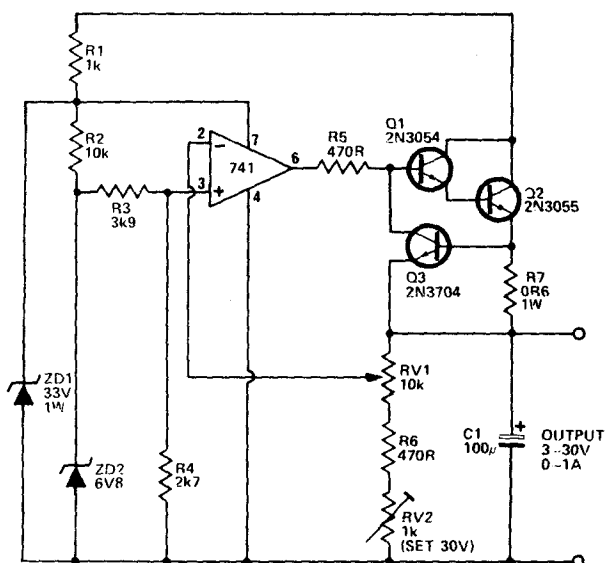
Fig. 26a Precision unity-gain half-wave rectifier.

act as a precision half-wave a.c./d.c. converter by designing it to give a voltage gain of 2.22 to give form-factor correction, and by integrating its rectifier output, as shown in Fig. 26b.

Note that each of the Fig. 26 circuits has a high output impedance, and the outputs must both be fed into loads having impedances less than about 1M.

Fig. 25 3V - 30V stabilised p.s.u. with overload protection.
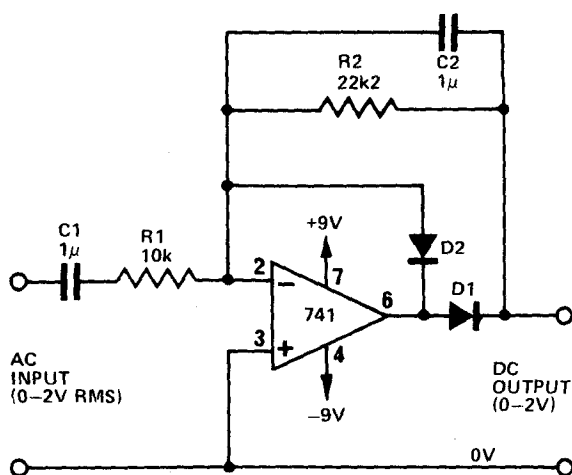
Fig. 26b Precision half-wave a.c./d.c. converter.

FIG. 27 shows how op-amp can be used as a high-performance d.c. voltmeter converter, which can be used to convert any 1V f.s.d. meter with a sensitivity better than 1k/V into a voltmeter that can read any

741 COOKBOOK

value in the range 1mV to 10V f.s.d. at a sensitivity of 1M/V· The voltage range is determined by the R_1 value, and the table shows some suitable values for common voltage ranges.

FIG. 28 shows a simple circuit that can be used to convert a 1mA f.s.d. meter into a d.c. voltmeter with any f.s.d. value in the range 100mV to 1000V, or into a d.c. current meter with any f.s.d. value in the range 1uA to 1A. Suitable component values for different ranges are shown in the tables.

Fig. 27 High-performance d.c. voltmeter converter.

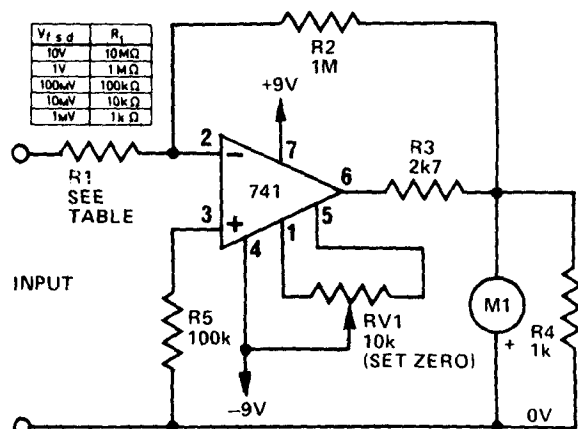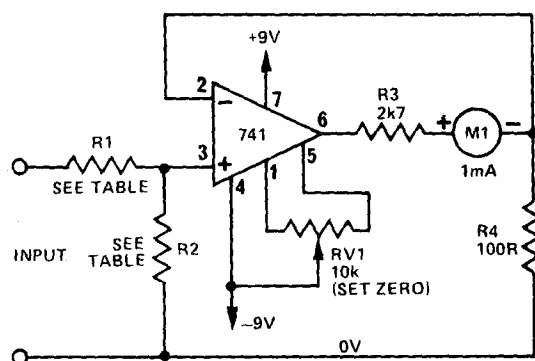| VOLTMETER			CURRENT METER		
f s d	R_1	R_2	f s d	R_1	R_2
1000V	10MΩ	1kΩ	1 A	0Ω	0 1Ω
100V	10MΩ	10kΩ	100mA	0Ω	1Ω
10V	10MΩ	100kΩ	10mA	0Ω	10Ω
1V	900kΩ	100kΩ	1mA	0Ω	100Ω
100mV	0Ω	100kΩ	100μA	0Ω	1kΩ
			10μA	0Ω	10kΩ
			1μA	0Ω	100kΩ

Fig. 28 Simple d.c. voltage or current meter.

FIG. 29 shows the circuit of a precision d.c. millivoltmeter, which uses a 1mA f.s.d. meter to read f.s.d. voltages from 1mV to 1000mV in seven switch-selected ranges.

FIG. 30 shows the basic circuit of a precision a.c. volt or millivolt meter. This circuit can be used with any moving-coil meter with a full scale current value in the range 100uA to 5mA, and can be made to give any full scale a.c. voltage reading in the range 1mV to 1000mV. The tables show the alternative values of R_1 and R_2 that must be used to satisfy different basic meter sensitivities, and the values of R_3 and R_4 that must be used for different f.s.d. voltage sensitivities.

HOME OHM

Finally, to conclude, Fig. 31 shows how the 741 op-amp can be used in conjuncton with a 1mA f.s.d. meter to make a linear-scale ohmmeter that has five decade ranges from 1k to 10M.

The circuit is divided into two parts, and consists of a voltage generator that is used to generate a standard test
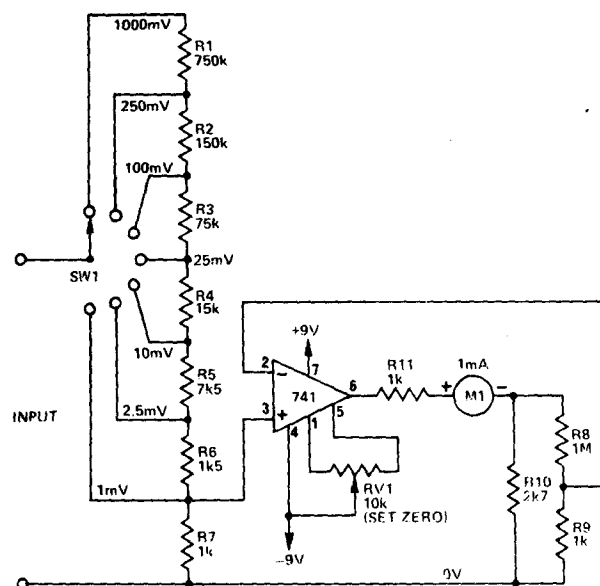
Fig. 29 Precision d.c. millivoltmeter.

M_1	R_1	R_2
100μA	9kΩ	27kΩ
500μA	1 8kΩ	5 6kΩ
1mA	900Ω	2 7kΩ
2 5mA	360Ω	1 5kΩ
5mA	180Ω	470Ω

VALUES FOR USE WITH DIFFERENT METER MOVEMENTS

$V_{f s d}$	R_4	R_3
1000V	10MΩ	10kΩ
100V	10MΩ	100kΩ
10V	10MΩ	1MΩ
1V	1MΩ	1MΩ
100mV	100kΩ	1MΩ
10mV	10kΩ	1MΩ
1mV	1kΩ	1MΩ

DIFFERENT f s d VOLTAGE SENSITIVITIES

Fig. 30 Precision a.c. volt/millivolt meter.

741 COOKBOOK

Fig. 31 Linear-scale ohmmeter.

voltage, and a readout unit which indicates the value of the resistor under test.

The voltage generator section of the circuit comprises zener diode ZD_1, transistor Q_1, and resistors R_1 to R_4. The action of these components is such that a stable reference potential of 1V is developed across R_4, but is adjustable over a limited range via RV_1. This voltage is fed to the input of the op-amp readout unit. The op-amp is wired as an inverting d.c. amplifier, with the 1mA meter and RV_3 forming a 1V f.s.d. meter across its output, and with the op-amp gain determined by the values of ranging resistors R_5 to R_9 and by negative feedback resistor R_x.

Since the input to the amplifier is fixed at 1V, the output voltage reading of the meter is directly proportional to the value of R_x, and equals full scale when R_x and the ranging resistor values are equal. Consequently, the circuit functions as a linear-scale ohmmeter.

CALIBRATION

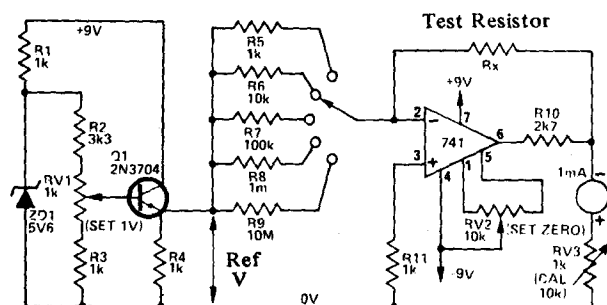The procedure for initially calibrating the Fig. 31 circuit is as follows: First, switch the unit to 10k range and fix an accurate 10kΩ resistor in the R_x position. Now adjust RV_1 to give an accurate 1V across R_4, and then adjust RV_2 to give a precise full scale reading on the meter. All adjustments are then complete, and the circuit is ready for use.

MISCELLANEOUS 741 PROJECTS

The 741 op-amp can be used as the basis of a vast range of miscellaneous projects, including oscillators and sensing circuits. Four such projects are described in this final section.

FIG. 32 shows how the 741 op-amp can be connected as a variable-frequency wien-bridge oscillator, which covers the basic range 150Hz to 1.5kHz, and uses a low-current lamp for amplitude stabilisation. The output amplitude of the oscillator is variable via RV_4 and has a typical maximum value of 2.5V r.m.s. and a t.h.d. value of 0.1%. The frequency range of the circuit is inversely proportional to the C_1-C_2 values· The circuit can give a useful performance up to a maximum frequency of about 25kHz.
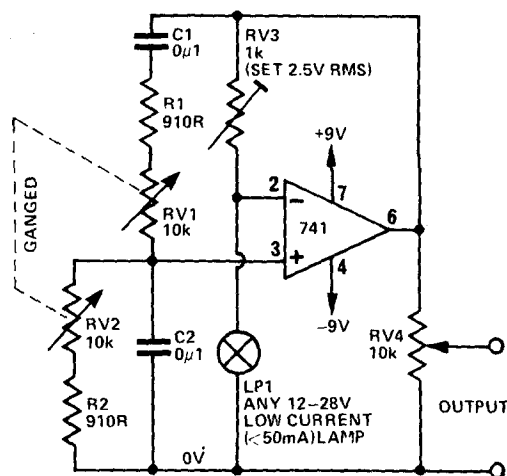

Fig. 32 150Hz - 1.5kHz Wien-bridge oscillator.

Fig. 33 shows how either a 741 or a 709 op-amp can be connected as a simple variable-frequency square-wave generator that covers the range 500Hz to 5kHz via a single variable resistor. (The circuit produces a good symmetrical waveform.)

The frequency of oscillation is inversely proportional to the C_1 value, and can be reduced by increasing the C_1 value, or vice-versa. The amplitude of the square wave output signal can be made variable, if required, by wiring a 10kΩ variable potential divider across the output terminals of the circuit and taking the output from between the pot slider and the zero volts line.
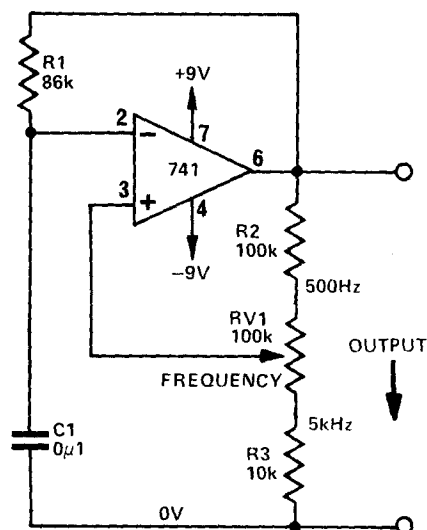
Fig. 33 Simple 500Hz - 5kHz square wave generator.

FIGS. 34 and 35 show a couple of useful ways of using the 741 op-amp in the open-loop differential voltage comparator mode. In each case, the circuits are powered from single-ended 12V supplies, and have a fixed half-supply reference voltage applied to the non-inverting op-amp terminal via the R_1-R_2 potential divider and have a variable voltage applied to the inverting op-amp terminal via a variable potential divider.

The circuit action is such that the op-amp output is driven to negative saturation (and the relay is driven on) when the variable input voltage is greater than the reference voltage. Conversely, the op-amp output is driven to positive saturation (and the relay is cut off) when the variable input voltage is less than the reference voltage.

741 COOKBOOK
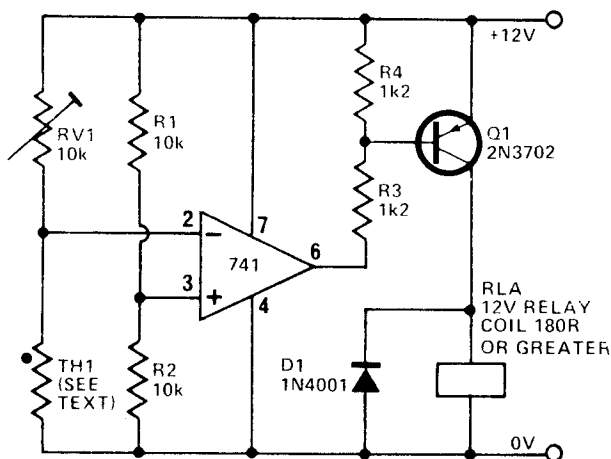
Fig. 34 Precision frost or under temperature switch can be made to act as a fire or over temperature switch by transposing R₁ and TH₁ positions.

FROSTY RECEPTION

The Fig. 34 circuit is that of a precision frost or under-temperature switch, which drives the relay on when the temperature sensed by thermistor TH_1 falls below a value pre-set via RV_1. The circuit action can be reversed, so that it operates as a fire or over-temperature switch, by simply transposing the RV_1 and the TH, positions. In either case, TH_1 can be any negative-temperature-coefficient thermistor that presents a resistance in the range 900Ω to $9k\Omega$ at the required trip temperature.
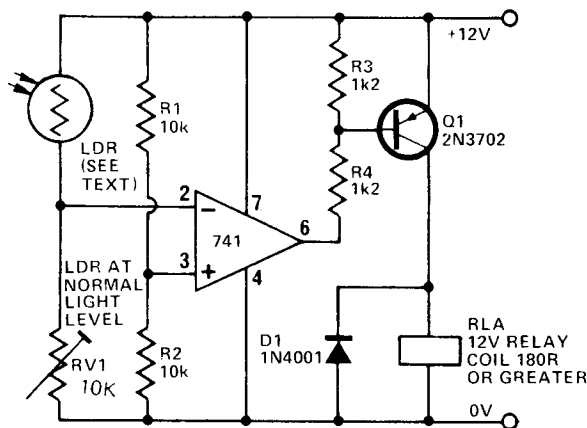
Fig. 35 Precision light-activated switch can be made to act as a dark-activated switch by transposing R₁ and LDR positions.

LIGHT WORK

The Fig. 35 circuit is that of a precision light-activated switch, which turns the relay on when the illumination level sensed by light-dependent resistor LDR exceeds a value pre-set by RV_1. The circuit action can be reversed so that the relay turns on when the illumination falls below a pre-set level by simply transposing the RV_1 and LDR positions. In either case, the LDR can be any cadmium-sulphide photocell that presents a resistance in the range 900Ω to $9k\Omega$ at the desired switch-on level.●

Lab Notes

Using the 3080 IC

This remarkable IC is quite different to run-of-the-mill op-amps as it has a "control" pin that varies the device's transconductance as the current into this input is varied. These circuits illustrate various ways to use the device — from an original manuscript by UK correspondent, Tim Orr.

THE CA3080 IS KNOWN as an *operational transconductance amplifier* (OTA). This is a type of op-amp, the gain of which can be varied by means of a control current, (I_{ABC}). The device has a differential input, a control input known as the 'amplifier bias input' and a current output. It differs in many respects from conventional op-amps and it is these differences that can be used to realise many useful circuit blocks.

Voltage controlled amplifier

The CA3080 can be used as a gain controlling device. A useful circuit is shown in Figure 1. The input signal is attenuated by R1, R2 such that a 20 mV peak-to-peak signal is applied to the input terminals. If this voltage is much larger, then significant distortion will occur at the output. In fact, this distortion is put to good use in the triangle-to-sinewave converter. (Figure 3, but we're jumping the gun).

The gain of the circuit is controlled by the magnitude of the current I_{ABC}. This current flows into the CA3080 at pin 5, which is held at one diode voltage drop above the $-$Vcc rail. If you connect pin 5 to 0 V, then this diode will get zapped (and so will the IC!). The maximum value of I_{ABC} permitted is 1 mA and the device is 'linear' over four decades of this current. That is, the gain of the CA 3080 is 'linearly' proportional to the magnitude of the I_{ABC} current over a range of 0.1 uA to 1 mA. Thus, by controlling I_{ABC}, we can control the signal level at the output.

The output is a current output which has to be 'dumped' into a resistive load (R5) to produce a voltage output. The output impedance seen at IC1 pin 6 is 10k (R5), but this is 'unloaded' by the voltage follower (IC2) to produce a low output impedance.

The circuit involving IC3 is a precision voltage-to-current converter and this can be used to generate I_{ABC}. When Vin (control) is positive, it linearly controls the gain of the circuit. When it is negative, I_{ABC} is zero and so the gain is zero.

This type of circuit is known by several names. It is a *voltage controlled amplifier*, (VCA), or an *amplitude modulator*, or a *two quadrant multiplier*.

One problem that occurs with the CA3080 is that of the 'input offset voltage'. This is a small voltage diffe-
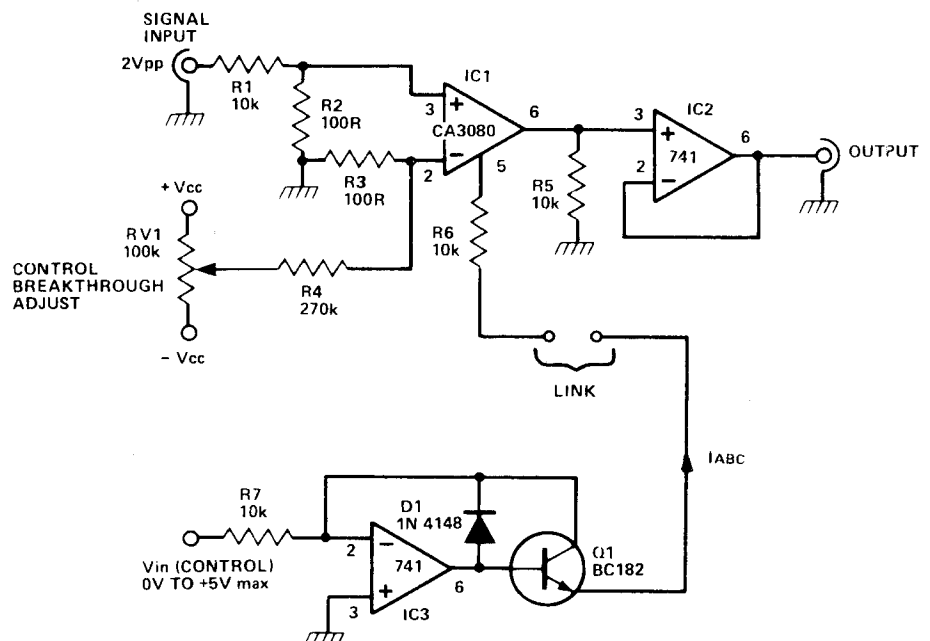
Figure 1. **A voltage controlled amplifier. Gain is varied by varying RV1. You can modulate a signal passing through the amplifier by joining the 'link' and applying a modulating signal to the input of IC3 (at R7). This sort of circuit is also known as a 'two quadrant multiplier'.**
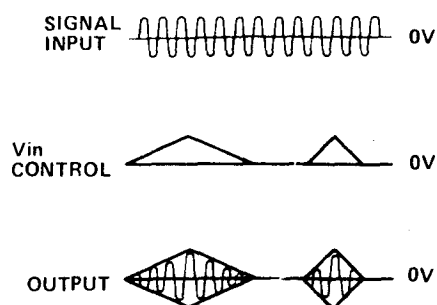
Figure 2. Illustrating the operation of the voltage controlled amplifier shown in Figure 1.

rence, or 'offset', between its input terminals. When there is no signal input and the control input is varied, a voltage similar to the control input will appear at the output. By adjusting RV1 it is possible to null out most of this control breakthrough.

The effect of modulating Vin (control) is illustrated in Figure 2.

Triangle to sinewave converter

By overloading the input of a CA3080 it is possible to produce a 'sinusoidal' transfer function. That is, if a triangle waveform of the correct magnitude is applied to the CA3080 input, the output will be distorted in such a way as to produce a sinewave approximation.
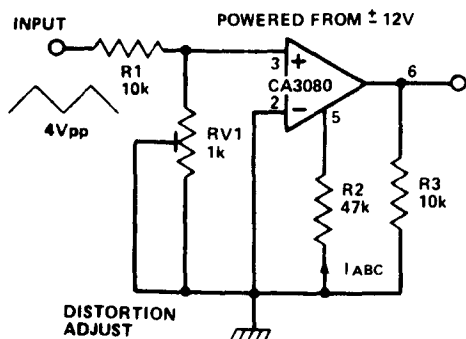
Figure 3. This circuit will convert a triangle wave to a sinewave with a resultant distortion of around 1.8%.

In the circuit shown (Figure 3), RV1 is adjusted so that the output waveform resembles a sinewave. I tested this circuit using an automatic distortion analyser and found the sinewave distortion to be only 1.8%, mostly third harmonic distortion which, for such a simple arrangement, seems very reasonable indeed. This could be used to produce a sinewave output from a triangle/square wave oscillator.
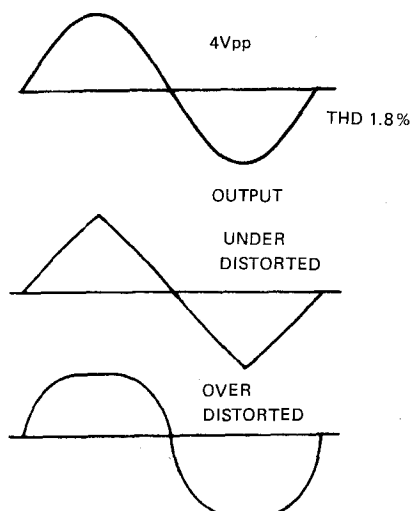
Figure 4. The output of the Figure 3 circuit should be adjusted (by RV1) to produce the waveform shown at top.
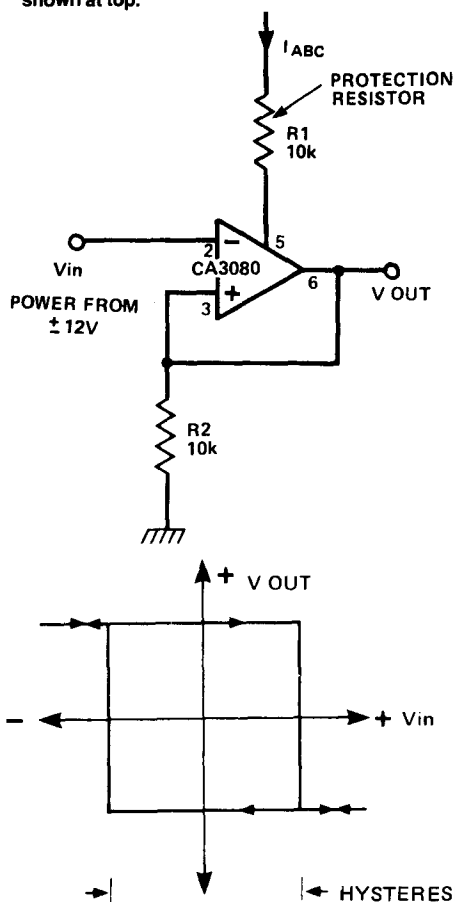
Figure 6. This sort of Schmitt trigger is not only simple but you can specify the hysteresis levels as well!
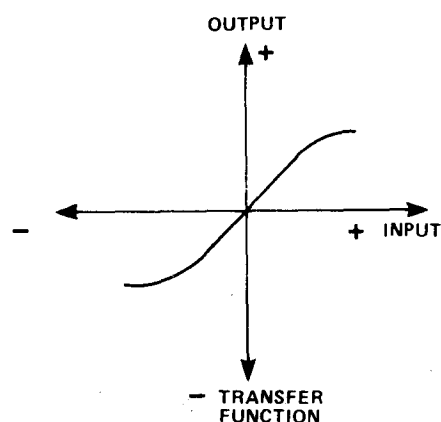
Figure 5. Transfer function of the Figure 3 circuit.
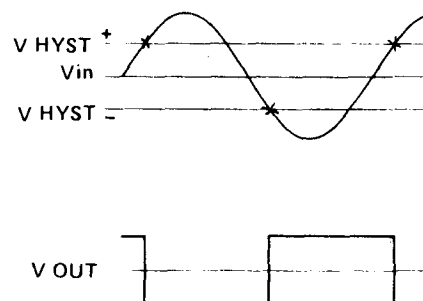
Figure 7. How the Schmitt trigger of Figure 6 works.

The result of varying RV1 is illustrated in Figure 4 and the transfer function of the circuit is shown in Figure 5.

Schmitt trigger

Most Schmitt trigger circuits prove to be very complicated when it comes to calculating the hysteresis levels. However, by using the CA3080 these calculations are rendered trivial, plus there is the added bonus of fast operation. The hysteresis levels are calculated from the simple equation,

$$V_{HYST} = +/- (I_{ABC} \times R2)$$

The output squarewave level is in fact equal in magnitude to the hysteresis levels. The circuit operation is as follows (referring to Figure 7):

Imagine the output voltage is high. The output voltage will then be equal to $(R2 \times I_{ABC})$ which we will call $+ V_{HYST}$. If V_{IN} becomes more positive than $+V_{HYST}$, the output will start to move in a negative direction, which will increase the voltage between the input terminals which will further accelerate the speed of the output movement. This ▶

31

SELECTED DATA ON THE 3080

Transconductance

TRANSCONDUCTANCE (gm) – (μmho) vs I_{ABC}–AMPLIFIER BIAS CURRENT (μA)

$V_S = \pm 15$ V; –55°C, +25°C, +125°C
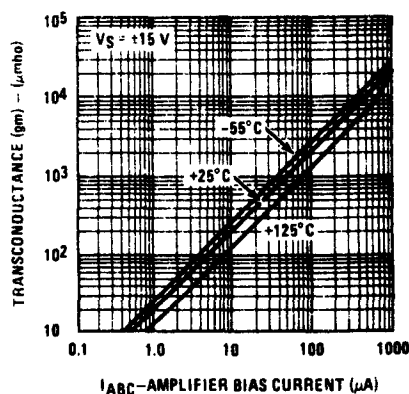
Input and Output Capacitance
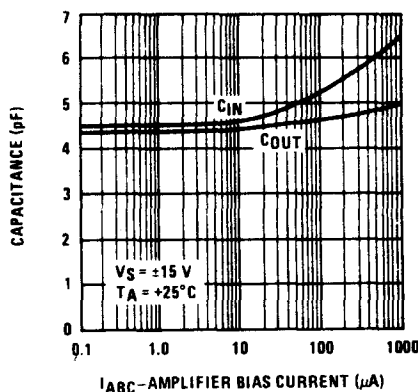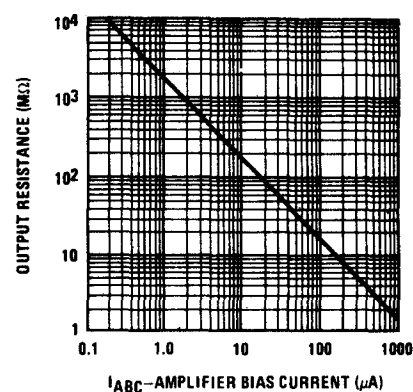
CAPACITANCE (pF) vs I_{ABC}–AMPLIFIER BIAS CURRENT (μA)

C_{IN}, C_{OUT}; $V_S = \pm 15$ V, $T_A = +25°C$

Output Resistance

OUTPUT RESISTANCE (MΩ) vs I_{ABC}–AMPLIFIER BIAS CURRENT (μA)

Peak Output Voltage and Common Mode Range

PEAK OUTPUT VOLTAGE AND COMMON MODE RANGE (V) vs I_{ABC}–AMPLIFIER BIAS CURRENT (μA)

V_{OUT}, V_{CMR}; $V_S = \pm 15$ V, $R_{LOAD} = \infty$, $T_A = +25°C$

Input Resistance

INPUT RESISTANCE (MΩ) vs I_{ABC}–AMPLIFIER BIAS CURRENT (μA)

Amplifier Supply Current

AMPLIFIER SUPPLY CURRENT (μA) vs I_{ABC}–AMPLIFIER BIAS CURRENT (μA)

$V_S = \pm 15$ V; +125°C, +25°C, –55°C
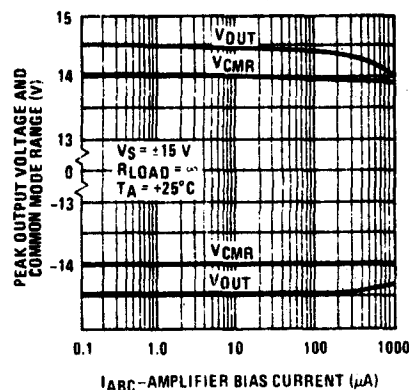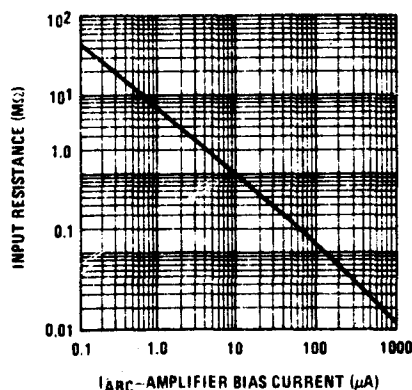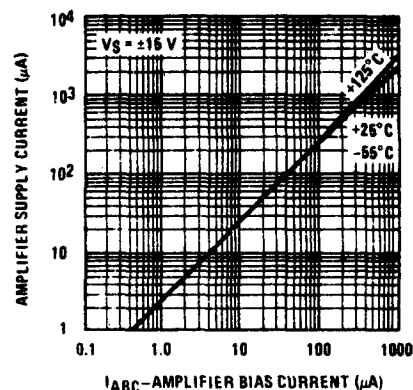
General description

The 3080 is a programmable transconductance block intended to fulfil a wide variety of variable gain applications. The 3080 has differential inputs and high impedance push-pull outputs. The device has high input impedance and its transconductance (gm) is directly proportional to the amplifier bias current (I_{ABC}).

High slew rate together with programmable gain make the 3080 an ideal choice for variable gain applications such as sample and hold, multiplexing, filtering, and multiplying.

Features

- Slew rate (unity gain compensated): 50 V/us
- Fully adjustable gain: 0 to gm R_L limit
- Extended gm linearity
- Flexible supply voltage range: +/–2 V to +/–18 V
- Adjustable power consumption

Absolute maximum ratings

Supply Voltage 3080	+/–18 V
Power Dissipation	250 mW
Differential Input Voltage	+/–5 V
Amplifier Bias Current (I_{ABC})	2 mA
DC Input Voltage	$+/V_S$ to $-V_S$
Output Short Circuit Duration	Indefinite

Electrical characteristics, 3080 (Note 1).

Parameter	Conditions	Min.	Typ.	Max.	Units
Forward Transconductance (gm)	Over Specified Temp. Range	6700	9600	13000	umho
		5400			umho
Peak Output Current	$R_L = 0$, $I_{ABC} = 5uA$		5		uA
	$R_L = 0$	350	500	650	uA
	$R_L = 0$	300			uA
	Over Specified Temp.Range				
Peak Output Voltage					
Positive	$R_L =$, 5uA I_{ABC} 500uA	+12	+14.2		V
Negative	$R_L =$, 5uA I_{ABC} 500uA	–12	–14.4		V
Amplifier Supply Current			1.1		mA
Common Mode Rejection Ratio		80	110		dB
Common Mode Range		+/–12	+/–14		V
Input Resistance		10	26		k
Open Loop Bandwidth			2		MHz
Slew Rate	Unity Gain Compensated		50		V/us

Note 1: These specifications apply for $V_S = +/–15$ V and $T_A = 25°C$, amplifier bias current (I_{ABC}) = 500uA, unless otherwise specified.

Internal circuit of the 3080

Internal circuit showing transistors T1–T9, diodes D1–D4, +V (pin 7), OUTPUT (pin 6), (–) INPUT (pin 2), (+) INPUT (pin 3), AMP BIAS INPUT (pin 5), –V (pin 4).

Figure 8. **This voltage controlled oscillator has a square and a triangle output. By varying the input between −5 V and +5 V, the output frequency can be varied over the range 10 Hz to 10 kHz. Note that Q1 and Q2 should be a matched pair.**

is known as regenerative feedback and is responsible for the Schmitt trigger action. The output snaps into a negative state at a voltage equal to $-(R2 \times I_{ABC})$ which is designated as $-V_{HYST}$. Only when V_{IN} becomes more negative than $-V_{HYST}$ will the output change back to the $+V_{HYST}$ state.

The Schmitt trigger is a very useful building block for detecting two discrete voltage levels and finds many uses in circuit designs.

Voltage controlled oscillator

By using two CA3080s and some 741 op-amps it is possible to make an oscillator, the frequency of which is voltage controllable. This unit finds many applications in the fields of electronic music production and test equipment.

The circuit (Figure 8) has been given a logarithmic control law, that is, the frequency of operation doubles for every volt increase in the control voltage. This makes it ideal for musical applications where linear control voltages need to be converted into musical intervals (which are logarithmically spaced) and also for audio testing where frequencies are generally measured as logarithmic functions.

One CA3080, IC2, is an integrator. The I_{ABC} current that drives this IC is used to either charge or discharge C1. This produces triangular waveforms which are buffered by IC3, which then drives the Schmitt trigger IC4. The hysteresis levels for this device are fixed at $+/-1.5$ V, being determined by R6 and R7.

The output of the Schmitt trigger is fed back in such a way as to control the direction of motion of the integrator's output. If the Schmitt output is high, then the integrator will ramp upwards and vice versa.

Imagine that the integrator is ramping *upwards*. When the integrator's output reaches the *upper* hysteresis level, the Schmitt will flip into its *low* state, and the integrator will start to ramp downwards. When it reaches the *low* hysteresis level the Schmitt will flip back into its *high* state. Thus the integrator ramps up and down in between the two hysteresis levels.

The speed at which it does this, and hence the oscillating frequency, is determined by the value of I_{ABC} for IC2.

The larger the current, the faster the capacitor is charged and discharged.

Two outputs are produced, a triangle wave (buffered) from IC3 and a square-wave (unbuffered) from IC4. If the squarewave output is loaded, then the oscillation frequency will change so a buffer is advisable. ▶
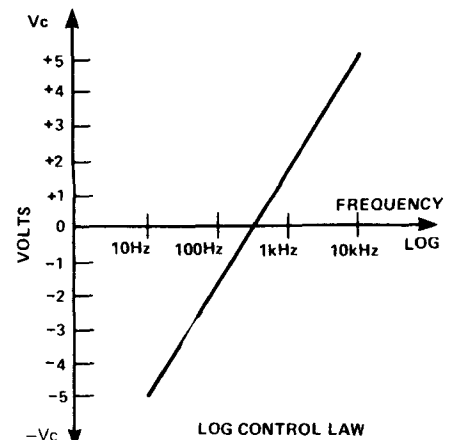

Figure 9. **Voltage versus frequency characteristic of the Figure 9 circuit.**

Lab Notes

The log. law generator is composed of Q1, 2, 3 and IC1. Transistors Q1 and Q2 should be matched so that their base emitter voltages (Vbe) are the same for the same emitter current, (50uA). Matching these devices to within 5 mV is satisfactory, although unmatched pairs could be used. When matching transistors, take care not to touch them with your fingers. This will heat them up and produce erroneous measurements.

Transistor Q2 is used to produce a reference voltage of about -0.6 V, which is connected to IC1 pin 3. This op-amp and Q3 is used to keep the emitter of Q1 at the same voltage of -0.6 V. The input control voltage is attenuated by R1, R2 such that a $+1$ V increase at the input produces a change of only $+18$ mV at the base of Q1. However, the emitter of Q1 is fixed at -0.6 V, so the current through Q1 doubles. (It is a property of transistors that the collector current doubles for every 18 mV increase in Vbe).

The emitter current of Q1 flows through Q3 and into IC2, thus controlling the oscillator frequency. It is possible to get a control range of over 1000 to 1 using this circuit. With the values shown, operation from 10 Hz to 10 kHz is achieved. Reducing C1 to 1n will increase the maximum frequency to 100 kHz, although the waveform quality may be somewhat degraded.

Changing C1 to 1uF, (non-polarised) will give a minimum frequency of 0.1 Hz.

Fast comparator

The high slew rate of the CA3080 makes it an excellent fast voltage comparator and a circuit is shown in Figure 10. When pin 2 of IC1 is more positive than Vref, the output of IC1 goes negative and vice versa. Vref can be moved around so that the point at which the output changes can be varied. As long as the input sinewave level is quite large (1 V say) then the output can be made to move at very fast rates indeed. However, care must be taken to avoid overloading the inputs. If the differential input voltage exceeds 5 V, then the input stage breaks down and may cause an undesired output to occur.

One use of a fast comparator is in a *tone burst generator*. A circuit is shown in Figure 11. This device produces bursts of sinewaves, the burst starting

and finishing on axis crossings of the sinusoid. The CA3080 is configured here as a voltage comparator, used to detect these axis crossings and to produce a square wave output which then drives a binary divider (IC3). The divider produces a 'divide by sixteen' output which is high for eight sinewave cycles and then low for the next eight. This signal is then used to gate ON and OFF the sinewave.

The gate mechanism is a pair of transistors which short the sinewave to ground when the divider output is high and let it pass when the divider output is low. The resulting output is a toneburst.

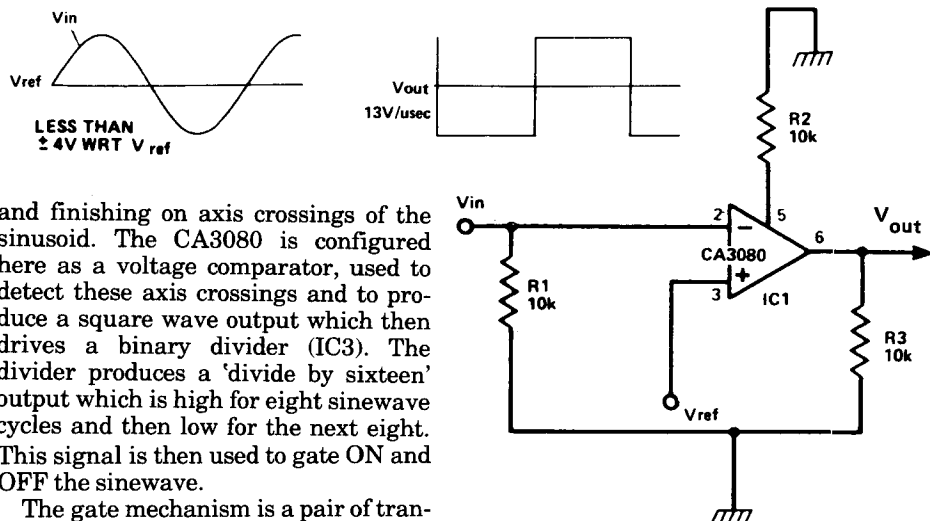However, if the comparator is not ▶

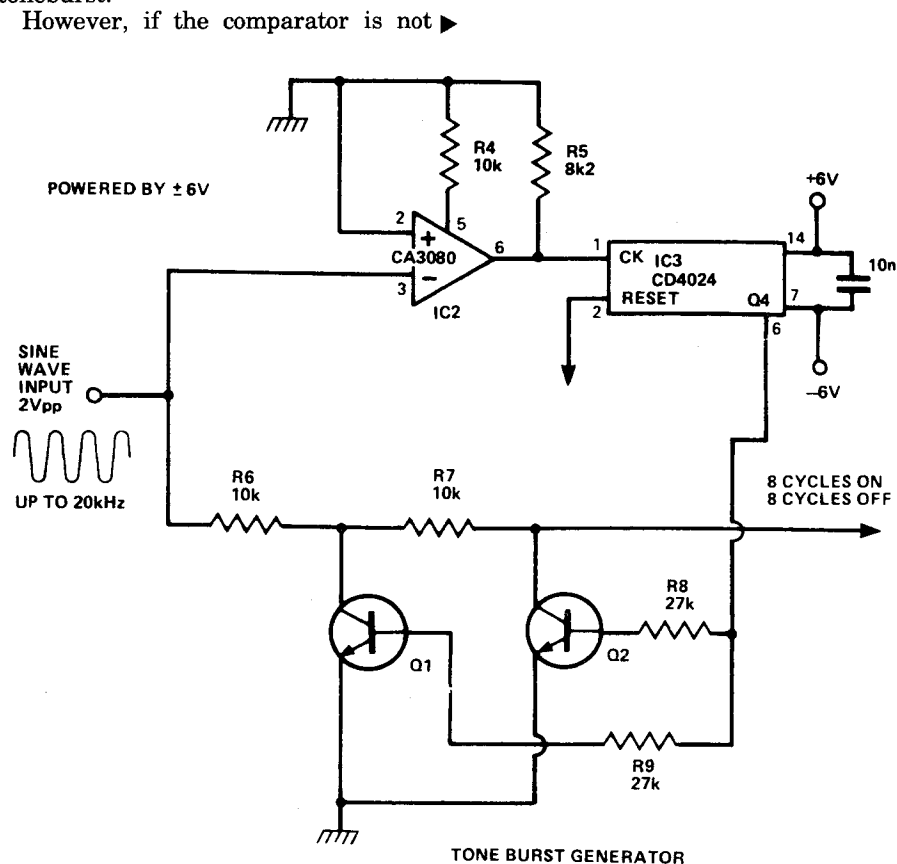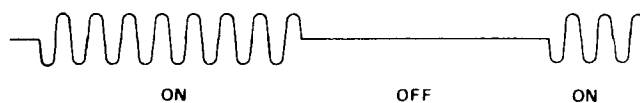Figure 10. **Example of a fast comparator.**

Figure 11. **A fast comparator is used in this tone burst generator, producing eight cycles of tone with eight cycle breaks starting and finishing at on-axis crossings.**
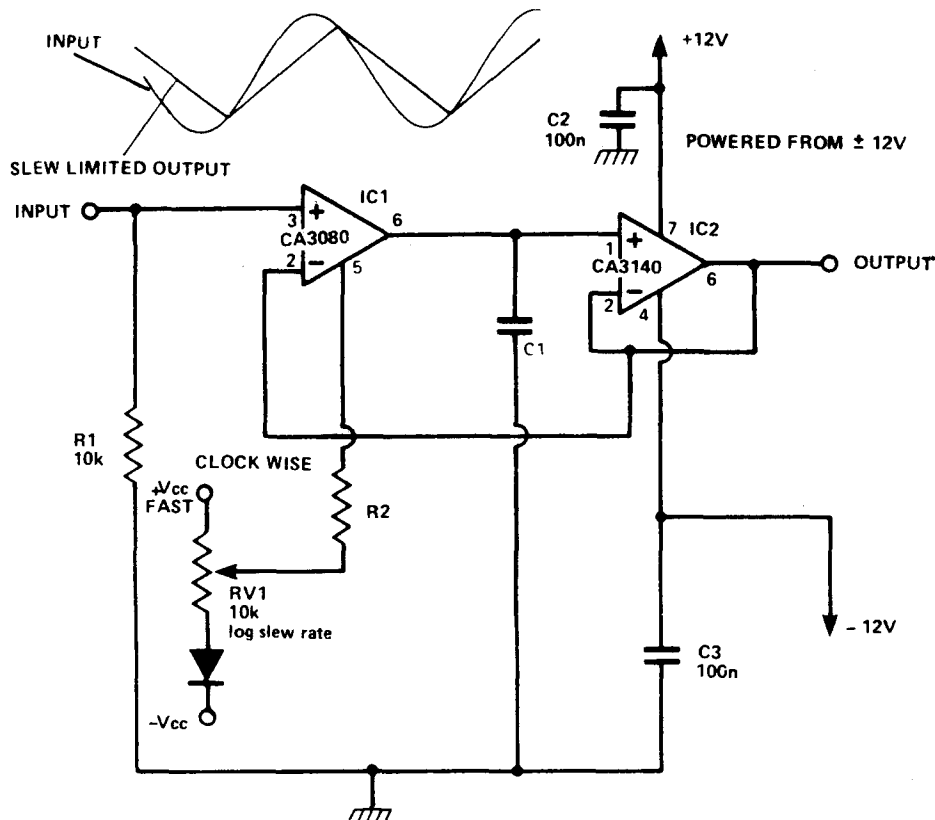
Figure 12. This slew rate limiter circuit produces a linear ramp on signals which exceed the slew rate limit, the output amplitude stopping when it reaches the signal level.
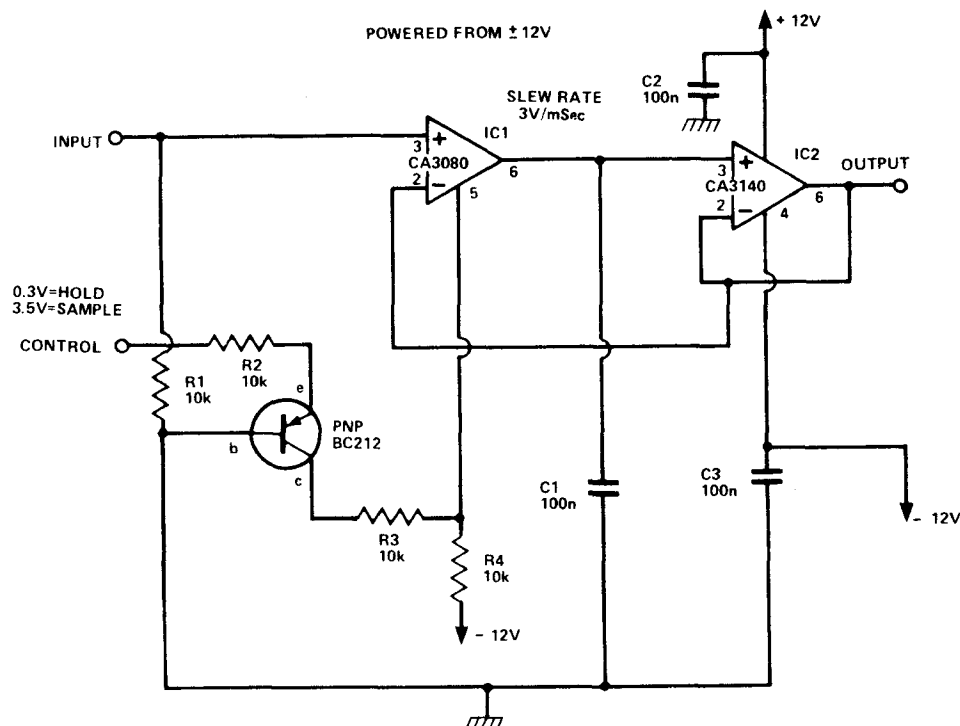
Figure 13. A typical application of the slew rate limiter is this sample and hold circuit.

very fast then there will be a delay in generating the gate and so the tone burst will not start or finish on axis crossings.

Using the circuit shown, operation up to 20 kHz is obtainable.

Slew limiter

The current output of a CA3080 can be used to produce a controlled slew limiter. By connecting the output current to a capacitor, the output voltage-cannot move faster than a rate given by

$$\text{Slew Rate} = \frac{I_{ABC}}{C1} \text{ Volts per sec.}$$

Note that I_{ABC} determines the slew rate and as I_{ABC} is a variable then so is the slew rate.

A suitable circuit is shown in Figure 12. The output voltage is buffered by a voltage follower, IC2. This is a MOSFET op-amp which has a very high input impedance, which is necessary to minimise the loading on C1.

When an input signal is applied to IC1 the output tries to move towards this voltage but its speed is limited by the slew rate. Thus, the output produces a linear ramp which stops when it reaches the input signal level.

Sample and hold

A typical application of the slew limiter circuit is in a *sample and hold* circuit. The circuit in Figure 13 could be termed an analogue memory. When the control voltage is high, the circuit will 're-member' or 'hold' the input voltage level present at the time. The result is shown in Figure 14. ▶

Figure 14. Illustrating the operation of the sample and hold circuit of Figure 13.

In this circuit, I_{ABC} is either hard ON (sample) or completely OFF (hold). In the sample mode, the output voltage quickly adjusts itself so that it equals the input voltage. This enables a short sample period to be used.

In the HOLD mode, I_{ABC} is zero and so the voltage on C1 should remain fixed.

Such circuits are used in music synthesizers (to remember the pitch), in analogue-to-digital converters and many other applications.

A multiplier/modulator

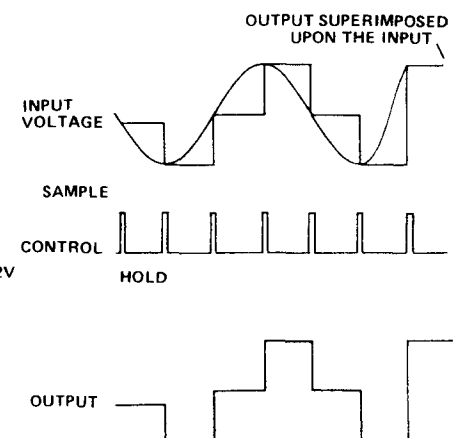The CA3080 is basically a two-quadrant multiplier, that is, it has two inputs, one of which can accept bipolar signals (positive and negative going) — the inverting or the non-inverting input — the other can only accept a unipolar signal — the control input, pin 5.

Whilst a two-quadrant multiplier is very useful in a wide variety of applications, a four-quadrant multiplier has extra advantages. For example, apart from amplitude modulation, it can perform frequency doubling and ring modulation. See Figure 16. Now, a four-quadrant multiplier has two inputs, both of which can accept bipolar signals. An example of a four-quadrant multiplier is a frequency converter in a radio receiver. The familiar diode ring mixer is another example of a four-quadrant multiplier.

The circuit in Figure 15 is fairly similar to that of the two-quadrant multiplier shown in Figure 1. This circuit has several important differences.
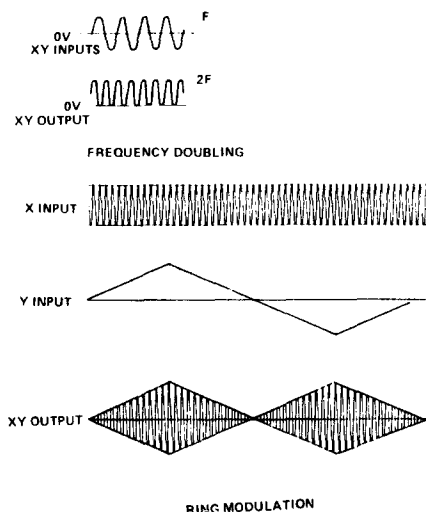

Figure 16. Illustrating the various operations of the four quadrant multiplier of Figure 15.
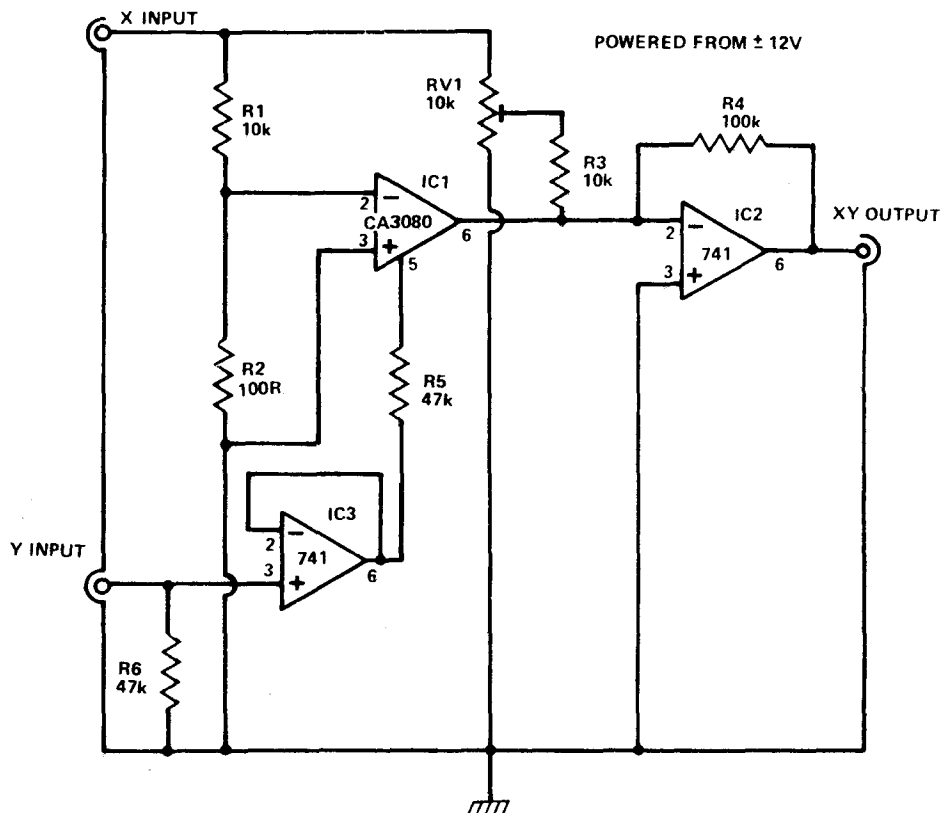

Figure 15. This multiplier/modulator can be used to produce a 'Dalek' voice when working as a ring modulator. It can also be used as a frequency doubler.

A 741 op-amp, IC3, is used to generate I_{ABC} in such a way that its input, the 'Y' input, can go both positive and negative. Thus, the Y input is bipolar.

When Y is at zero volts (no input) and there is a signal on the X input the desired output (X x Y) should be zero. This is achieved by adjusting RV1 so that the signal via IC1 (this is inverted) is exactly cancelled out by that via R3. Now, when Y is increased positively, a non-inverted value of X is produced at the output and, when Y is increased negatively, an inverted value of X is produced. When Y is zero, so is the output. This is known sometimes as ring modulation.

If a speech signal is connected to the X input and an audio oscillator to the Y input, the resulting sound is that of a 'Dalek'.

Also, if a sinewave is connected to both the X and Y inputs, the XY product is a sinewave of twice the frequency. This is known as a frequency doubler, but it will only work with sinewaves.

For more theoretical information on four-quadrant multipliers, especially the variable transconductance type, see "Operational Amplifiers" (second edition), by G.B. Clayton, published by Newnes-Butterworths and available in Australia through Butterworths, 586 Pacific Highway, Chatswood NSW 2067. (02)412-3444.

Single pole filter/wah wah

The guitar 'wah wah' effects unit employs a filter which can be manually 'swept' across the middle of the audio frequency range, generally from around 500 Hz to 5 kHz or so, producing the peculiar 'wah wah' sound.

A single pole, voltage-controlled, low pass filter can be constructed using a CA3080 as a current-controlled resistor. The circuit is shown in Figure 17.

A simple, low pass RC filter configuration is employed, the controllable 'R' is the CA3080 and the 'C' is C1. Varying I_{ABC} varies the amount of current drive to C1. This circuit configuration would normally be a slew limiter, except that the signal level to the input of the CA3080 is kept deliberately low (R1 and R2 form a 100:1 attenuator) so that the IC operates in its linear mode. This enables it to look like a variable resistor. ▶
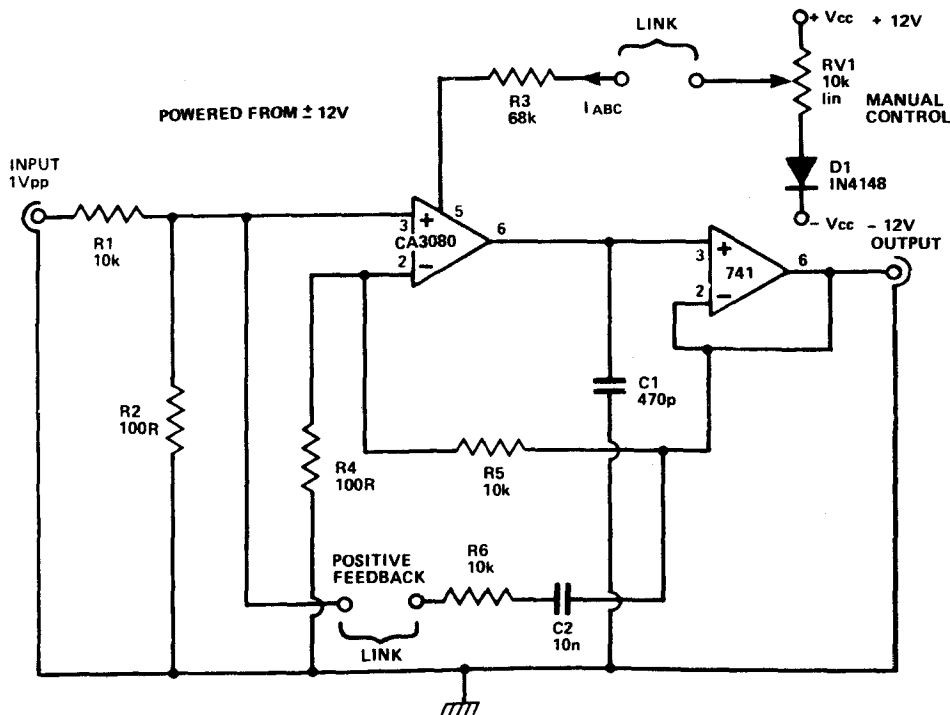
When this resistor is varied, the break frequency of the filter also varies.

By applying some positive feedback around the filter (R6, C2) it is possible to produce a peaky filter response. The peak actually increases with frequency, producing the wah wah effect.

The circuit as shown can be swept from about 400 Hz at the lower extreme to about 4 kHz at the upper extreme. See Figure 18.

Figure 17. A guitar wah-wah unit can be made with a swept frequency single pole filter.
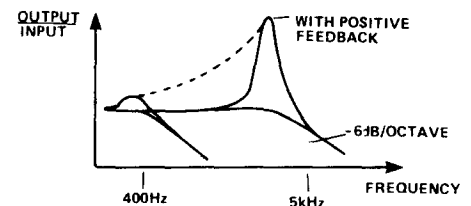
Figure 18. How the single pole filter affects the frequency response of the signal passed through the wah wah unit.
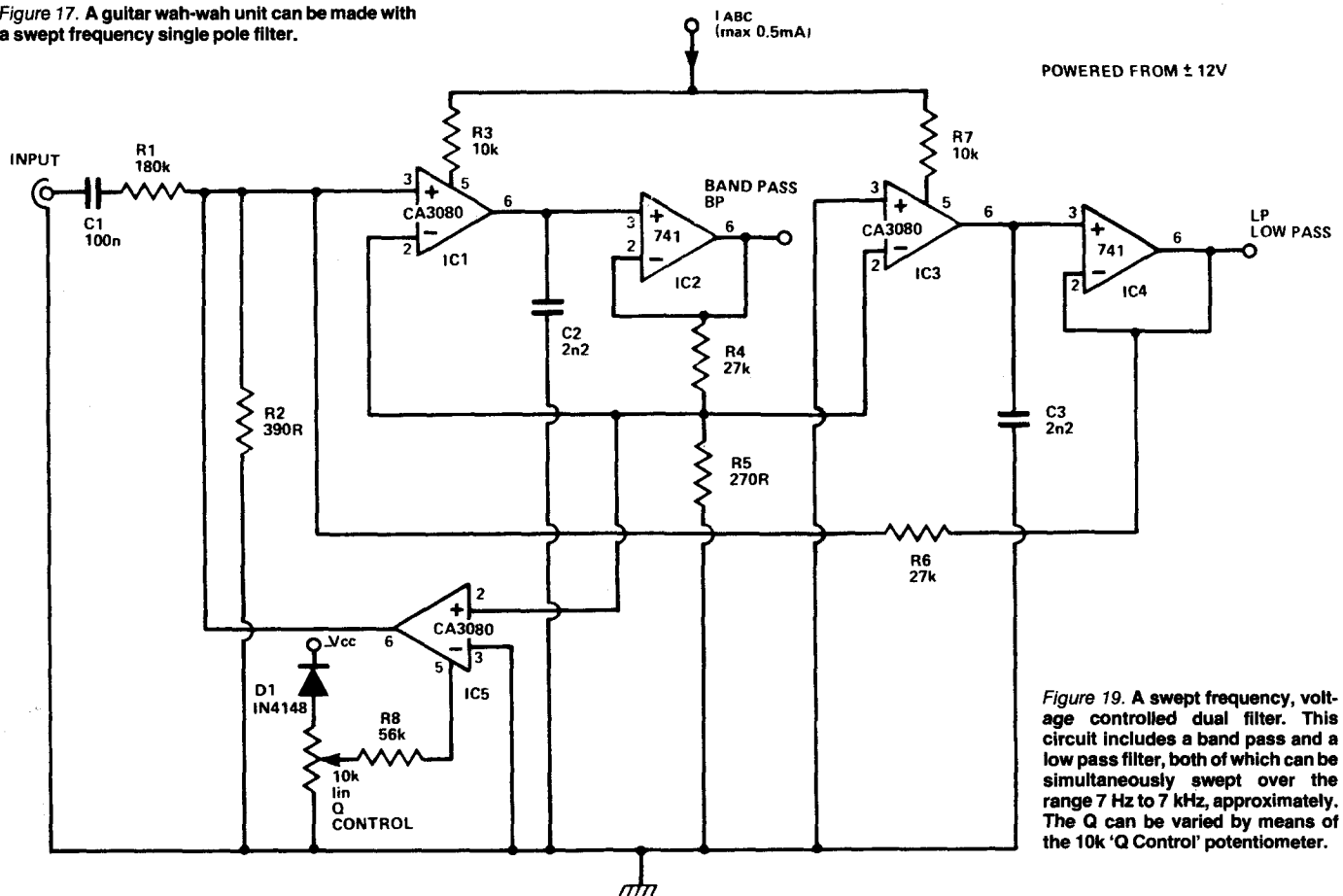
Figure 19. A swept frequency, voltage controlled dual filter. This circuit includes a band pass and a low pass filter, both of which can be simultaneously swept over the range 7 Hz to 7 kHz, approximately. The Q can be varied by means of the 10k 'Q Control' potentiometer.

37

Lab Notes

Voltage controlled filter

A standard dual integrator filter can be constructed using a few CA3080s. By varying I_{ABC}, the resonant frequency can be swept over a 1000 to 1 range. IC1 and IC3 are two current-controlled integrators. IC2 and IC4 are voltage followers which serve to buffer the high impedance outputs of the integrators. A third CA3080 (C5) is used to control the Q factor of the filter. Q factors as high as 50 can be obtained. The resonant frequency of the filter is linearly proportional to I_{ABC} and hence this unit is very useful in electronic music production.

There are two outputs, a low pass and a band pass response. Minimum frequency is around 7 Hz to 10 Hz, upper frequency is around 7 kHz or so. Changing C2 and C3 will alter the upper and lower frequency limits. ●

Figure 20. Illustrating the operation of the filters in the Figure 19 circuit.

Active Filter Cookbook

Design your next filter the easy way — this article by Tim Orr
shows you how.

THERE ARE THREE main types of filter — low-pass, band-pass, and high-pass. Each does more or less what its name implies. A low pass filter passes all frequencies below the so-called 'roll-off' point and increasingly blocks all frequencies above this point. A band-pass filter passes all frequencies above a lower 'roll-off' point and below a higher roll-off point. A high-pass filter passes all frequencies above the roll-off point.

Firstly, consider the simple low-pass filter shown in Fig. 1a. The frequency response (shown in Fig 1b) is nearly flat until the break point — shown as fb. Above this point the response rolls off at 6 dB/octave. The break point is defined as the frequency where the resistance equals the capacitive reactance. At this point the output is attentuated to 0.707 (−3 dB) of the input. Although the resistance equals the capacitive reactance, the output is not half of the input. It is the vector sum of the two and hence is 0.707 of the input.

As the frequency response is a complex curve it is commonly approximated by a straight line. Such a line is called an asymptote (Fig. 1c). Note the frequency response graph uses logarithmic scales, octave or decades along the frequency axis, and dBs along the vertical axis representing output voltage divided by input voltage.

Phase shift with respect to frequency is often plotted as in Fig. 1d. Phase and frequency response plots are also known as Bode diagrams and are most useful in showing a filter's performance.

Note that for the low-pass filter of Fig. 1a, phase shift starts at 0°, is 45° at fb and approaches 90° as frequency approaches infinity. This is not an active filter. It is made up from passive com-

ponents and its output cannot be loaded substantially without changing its performance.

Figure 1e shows the same filter in active form, the op amp being used as a voltage follower serving only to isolate the filter's output. This configuration is known as a first order filter —
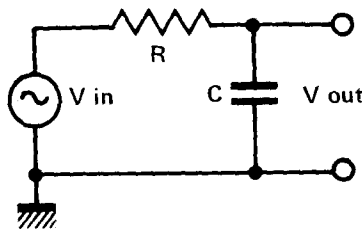
Fig. 1a. Simple low pass filter.

Fig. 1b. Frequency response

$$fb = \frac{1}{2 \pi CR}$$

WHERE fb IS IN Hz
C IS IN FARADS
R IS IN OHMS
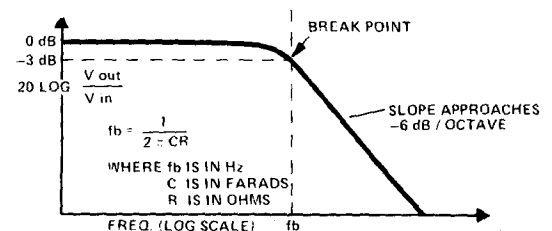
*NOTE −6 dB/OCTAVE = −20 dB/DECADE
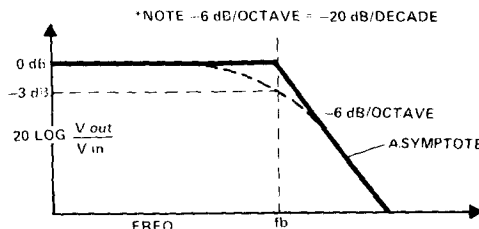
Fig. 1c. Approximation to response.
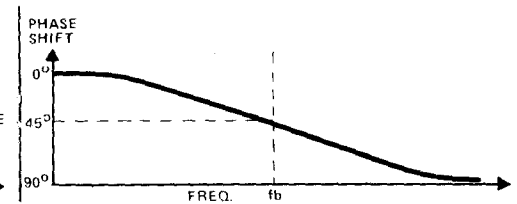
Fig. 1d. Phase shift v Frequency plot.
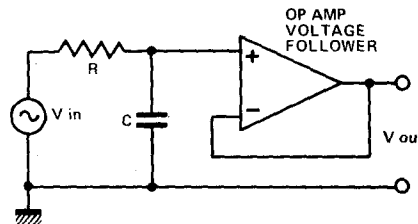
Fig. 1e. Active filter to perform
the same task as the passive
circuit of Fig. 1a.

Summary of low pass filter of Fig. 1

Filter type	Low pass
Filter order	First order
Roll off slope	−6dB/octave
Breakpoint fb	$fb = 1/2\pi CR$ Hz
Phase shift at fb	45°

Active Filter Cookbook

the expression 'first order' being an indication of the roll-off slope.

When a steeper slope is required, a higher order filter (that is, one with more elements) must be used. These are dealt with later.

Passing Highs

The simple high-pass filter shown in Fig 2a is the complement of the low-pass filter — the elements have simply been interchanged. Hence the complementary curves of Fig 2b. Note the break point and roll-off slope are similar.

Passing bands

A simple band-pass filter is shown in Fig 3a. Although it uses an inductor this is only to illustrate the band-pass theory.

The frequency response (Fig 3b) is symmetrical, rolling off at 6 dB/octave on either side of its peak. This filter is called a second order filter because it has two reactive sections (L and C). The C produces the +6 dB/octave portion of the slope, the L the -6 dB portion. The response of the filter peaks, and the slopes become much steeper where these two slopes meet.

The sharpness of the peak determines the quality of the filter (Q). Resonance occurs at the frequency known as the centre frequency — shown on our drawing as fc.
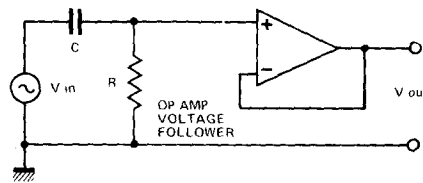
Fig. 2a. Simple high-pass active filter.

The band-pass filter is so-called because it passes signals within a certain bandwidth. This bandwidth is defined as being the frequency range contained between the two points that are 3 dB

Fig. 3c. Band-pass phase response.
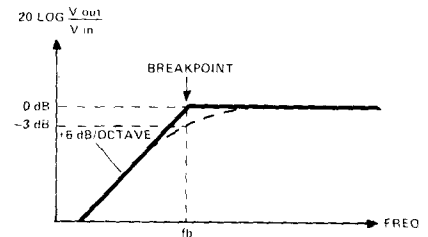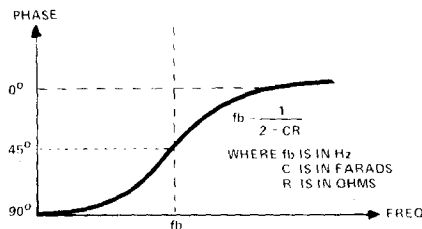
Fig. 2b. Frequency response (above) and phase response (left) of the high pass filter.

Summary of the high pass filter of Fig. 2.

Filter type	High pass
Filter order	First order
Roll off slope	+6dB/octave
Break point fb	$fb = 1/2\pi CR$ Hz
Phase shift at fb	45°

TABLE 2

below the resonant peak. Thus there is a fixed relationship between centre frequency (fc), bandwidth (fbw), and Q. The centre frequency is $fc = 1/2\pi\sqrt{LC}$.

CENTRE FREQUENCY fc

$$fc = \frac{1}{2\pi\sqrt{LC}}$$

WHERE fc IS IN Hz
L IS IN HENRYS
C IS IN FARADS

TUNED CIRCUIT

Fig. 3a. Simple band-pass filter.

Fig. 3b. Band-pass frequency response.
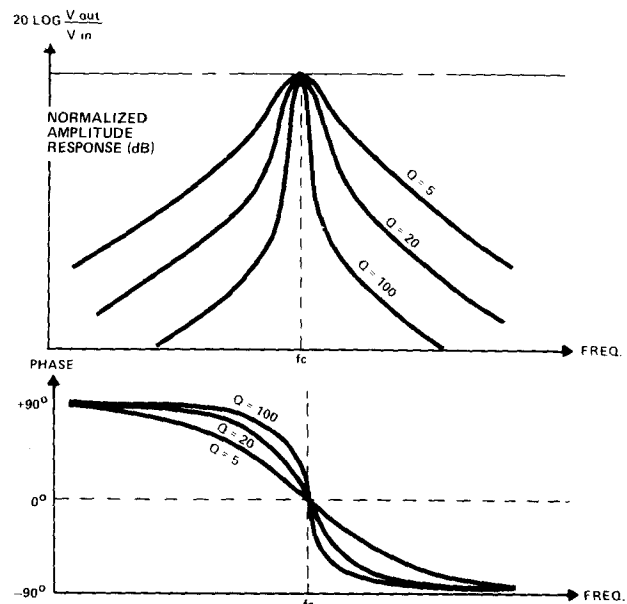
Fig. 3d. Effect of varying Q on the frequency and phase characteristics.

40

This is only approximate as it assumes that the value of R is relatively low. As R decreases, the Q increases. Thus R has the effect of damping the resonances, and as it approaches zero ohms, Q approaches infinity.

The phase shift is shown in Fig 3c. As this filter is a second order structure the total phase movement will be twice that of a first order structure, i.e. 180°. Figure 3d shows the phase and frequency responses for different values of Q. Note that a high Q has a very rapid rate of change of phase, a low Q has only a slow rate of change.

Time response

Band-pass filters have a time response as well as a frequency response. When an impulse is applied to a band-pass filter it rings (Fig 3e). The filter oscillates at the centre frequency (fc), the amplitude of oscillations decaying exponentially with time. The ringing time Tr is the time taken for the oscillations to decay to 37% of their initial value.

Ringing time is related to Q and fc by the following equation:—

$$Tr = Q/2\pi fc$$

In practice it may prove difficult accurately to measure the Q of a high-Q filter because the band-width is narrow. However if the filter can be made to ring a reasonably accurate measurement of Q can be obtained by measuring Tr and fc.

Notch filters

Another common type of filter is the band-reject or notch filter. There are many ways of building these, one way is shown in Fig 4. The input signal is subtracted from the band-pass output.
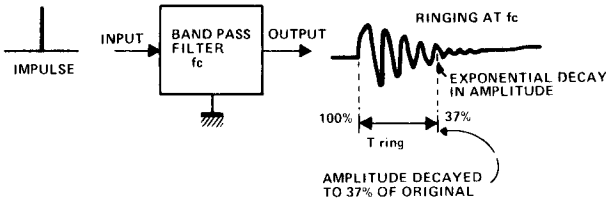
Fig. 3e. Ringing in a band-pass filter.

TABLE 3. Summary of band-pass filter.

Filter type	Band pass
Filter order	Second order
Roll off slopes	+ and —6dB/octave greater near to resonance
Centre frequency fc	$fc \sim 1/2\pi\sqrt{LC}$
Phase shift at fc	0
Q factor	fc/fbw where flow is the 3dB bandwidth
3dB bandwidth fbw	fc/Q
Ringing time, Tr	$Q/2\pi fc$
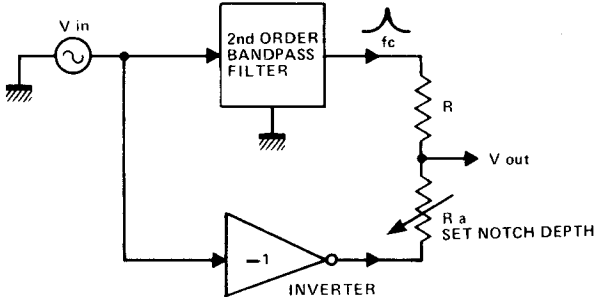
Fig. 4. Notch filter using Op-Amps.

By adjusting Ra with respect to R, complete cancellation can be obtained at fc. Thus the centre frequency of the band-pass filter is the centre frequency of the notch. The depth of this notch can be varied by altering the value of Ra. Very deep notches are possible: 50 dB being readily obtainable. As the Q of

the band-pass filter is increased so is the Q of the notch filter. Note that Ra must be adjusted for each value of Q.

Filter orders

Consider the ideal filter shown in Fig 5a. Its response is flat right up to the break frequency with frequencies above fb attenuated to zero. It looks really

Fig. 5a. Ideal low-pass response;
b,c. Examples of maximally flat filter responses;
d. Effect upon band-pass filter of increasing order number.

41
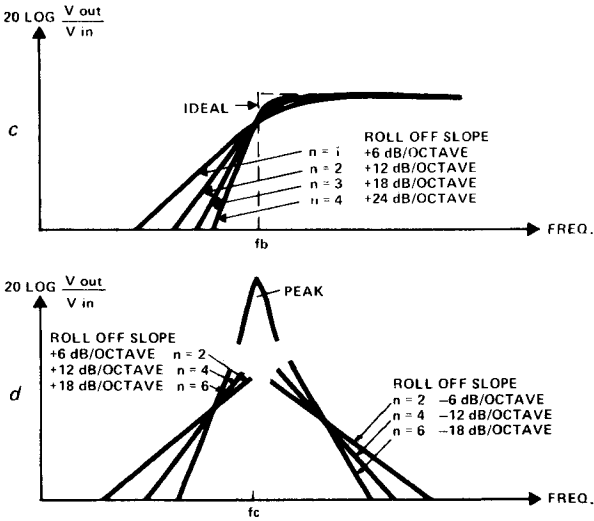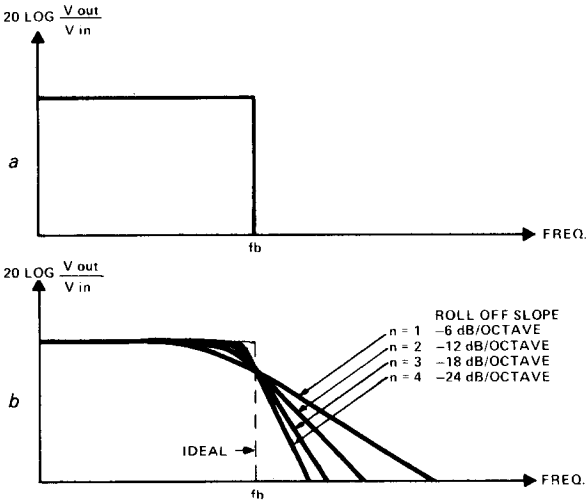
Active Filter Cookbook

good but in real life filters like this don't exist!

There is nevertheless a frequent need to design filters with very steep roll-off slopes. This is achieved by designing filters with lots of sections — thus increasing the filter order.

Each reactive element in a filter increases the filter order by one, therefore a low pass active filter with three capacitors is known as a third order filter and will have an ultimate roll-off of three times 6 dB/octave i.e. 18 dB/octave.

Unfortunately there is more to designing a third order low-pass filter than just sticking three first order circuits in line astern. That merely results in a very soggy curve! The filter should be flat in the pass band, then it should turn over and rapidly assume its roll-off slope. Examples of maximally flat filters are shown in Figs 5b and 5c. The effect of order number on a band-pass filter is shown in Fig 5d.

Later in this article circuit diagrams and design charts are shown for various filter types and order numbers. It might seem that to persuade a filter to approach its ideal response all that is needed is to increase the order number. This is of course so — but there are problems in obtaining components of sufficient accuracy. An eighth order filter for example needs components having values within 1% — possible with resistors but virtually impossible for capacitors.

Filter Shape

The type of filter required to do a certain job will depend on what parameters are most important. Three basic characteristics must be considered (high-pass and low-pass only).

1. Good transient response
2. Maximum flatness within the pass band
3. Steep roll-off slope.

Filters have been categorised into three basic types for simplicity.

BESSEL FILTER: phase changes almost linearly with frequency, useful for systems where a good transient response is required — such as joining all the little pulses on the output of a digital-to-analogue convertor. Very poor initial roll off.

BUTTERWORTH FILTER: This has the flattest pass band possible. Its two other parameters are a compromise — a reasonable overshoot and a fairly fast initial roll-off.

CHEBYSHEV FILTER: This has a small amount of ripple in its pass band, a very fast initial roll off but a poor transient response.

Rolling your own

In all the examples which follow the filters have been designed for operation at 1 kHz. To change the operating frequency resistor/s RF must be scaled accordingly (note: resistors RD are not changed).

For example if the filter is required to operate at 250 Hz then RF must be multiplied by $\frac{1000}{250}$. Figure 7 shows a first order low pass filter.

Figure 8a, b, and c shows second, third and fourth order filters.

The total design procedure is as follows:

1. Decide which type of filter is required — low, band-pass or high.
2. In the case of low or high-pass decide which type of response is required, Bessel, Butterworth of Chebyshev.
3. Decide what filter order is needed. This will lead you to a particular order filter with components shown scaled for 1 kHz.
4. Scale the resistors RF accordingly.
5. Build and test the filter.

As an example let us design an audio scratch filter having a break frequency of 7.5 kHz and an attentuation of more than 20 dB at 15 kHz.

The first decision is type of response required. A roll-off of more than 20 dB/octave is quite steep and so the Bessel filter is ruled out. The Chebyshev has poor transient response and we'd hear it ringing at 7.5 kHz. So we're left with the Butterworth.

Next comes filter order. Third order gives —18 dB/octave: this is not enough. Fourth order gives —24dB/octave. So a fourth order Butterworth filter it is.

Break frequency is 7.5 kHz so resistors RF and RF2 must be divided by 7.5. This gives the following rather funny values —
RF1 = 1k42, RF2— 1k42, RD1 = 5k9,

ATTENUATION AT FIRST OCTAVE (2 fb)

*3 dB CHEBYSHEV	17	28	39	51	62	75
BUTTERWORTH	12	18	24	30	36	42
1/FILTER ORDER	2	3	4	5	6	7

*NOTE THE IMPROVED ATTENUATION
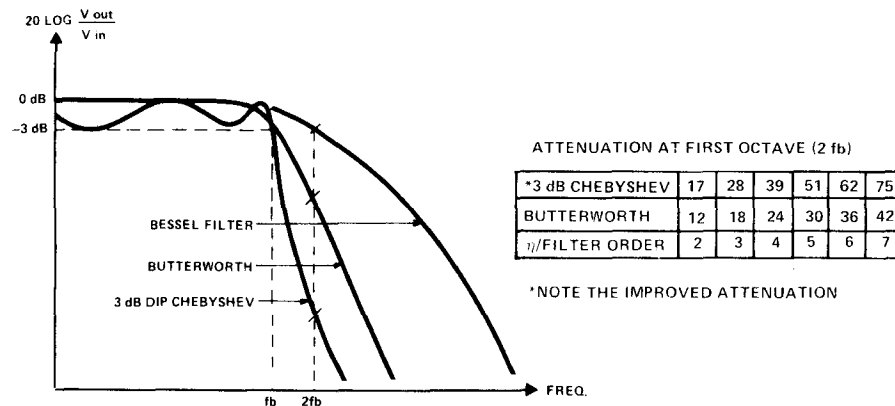
Fig. 6. Response of all three types of filter discussed, with table showing variation in attenuation between them.

*IN ALL LOW PASS FILTERS, THE SIGNAL SOURCE V in MUST PROVIDE A LOW IMPEDANCE DC PATH TO GROUND.

	RF	GAIN IN dB	COMPONENT TOLERANCES	
BESSEL	10k66	0	10%	ALL THE RESPONSES ARE
BUTTERWORTH	10k66	0	10%	THE SAME FOR A FIRST
CHEBYSHEV	10k66	0	10%	ORDER FILTER

Fig. 7. A general circuit for a first order low-pass filter.

RD2 = 48k7, C = 15 nF with component tolerances of 5%. We must now fit preferred values to our theoretically derived numbers.

Resistor RD2 can be 47k, RD1 6k2 (just over the limit of tolerance). Resistors RF1 and RF2 are a bit of a problem. The solution is to use the nearest 1% tolerance resistor or use 1K5. This will lower the break frequency by 6% but as this is an audio filter it won't matter too much.

Figure 9 provides design data for high-pass filters. The design procedure is exactly the same as for the low-pass designs.

A few problems may occur with the final results. One is that these filters have a voltage gain in their pass band — so you might find that although you have the required frequency response there is unexpected signal gain.

This may cause some problems with op amp band-width. As a rule of thumb, op amps should have 10 to 100 times more band-width than the product of the filters' maximum operating frequency multiplied by the individual stage gain of each section. If the op amp runs out of band-width or introduces a phase shift the filter won't work properly. For the examples given if you use a

ALL THE RESPONSES ARE THE SAME FOR A FIRST ORDER SYSTEM

	RF1	GAIN IN dB	COMPONENT TOLERANCE
BESSEL	10k66	0	10%
BUTTERWORTH	10k66	0	10%
CHEBYSHEV	10k66	0	10%

	RF1	RD1	GAIN IN dB	COMPONENT TOLERANCE
BESSEL	13k55	10k5	1.3	10%
BUTTERWORTH	10k66	22k6	1.6	10%
CHEBYSHEV	9k01	48k7	2.2	5%
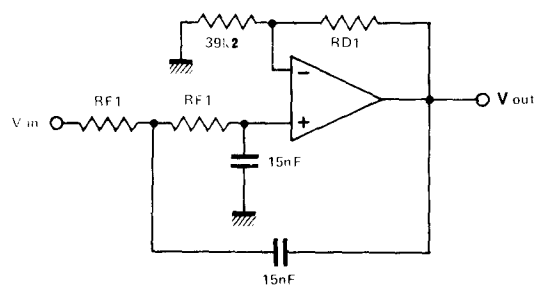
	RF1	RD1	GAIN IN dB	COMPONENT TOLERANCE
BESSEL	8k39	10k5	2.3	10%
BUTTERWORTH	10k66	2?k6	4.1	10%
CHEBYSHEV	12k6	48k7	6.8	5%

Fig. 8a. Second-order low-pass filter design, break frequency = 1 kHz.
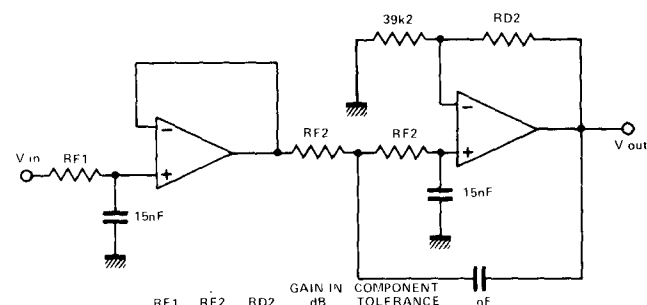
	RF1	RF2	RD2	GAIN IN dB	COMPONENT TOLERANCE
BESSEL	8k	7k26	21k5	4.1	10%
BUTTERWORTH	10k66	10k66	39k2	6.0	10%
CHEBYSHEV	35k41	11k73	66k5	8.6	2%

Fig. 8b. Third order low-pass filter. To alter break frequency (in this case 1 kHz) scale resistors accordingly.

	RF1	RF2	RD2	GAIN IN dB	COMPONENT TOLERANCE
BESSEL	14k19	15k68	21k5	4.1	10%
BUTTERWORTH	10k66	10k66	39k2	6.0	10%
CHEBYSHEV	3k21	9k70	66k5	8.6	2%

Fig. 9. From the top: First, second and third order high-pass filters, break point 1 kHz. Final roll-off is 6, 12, and 18 dB/octave respectively.

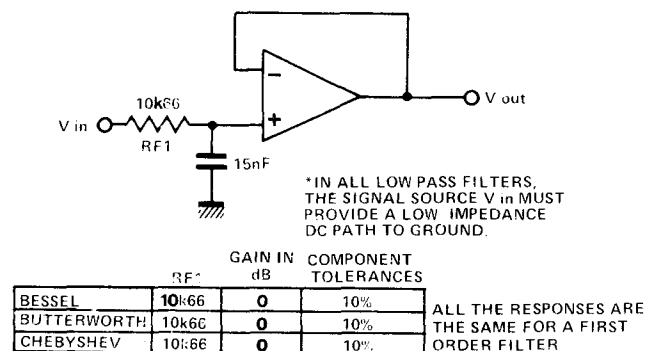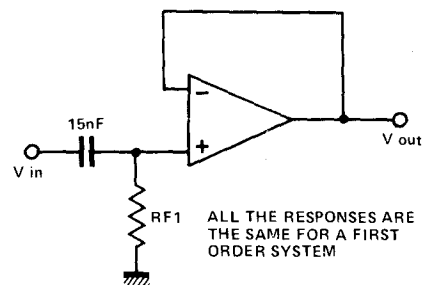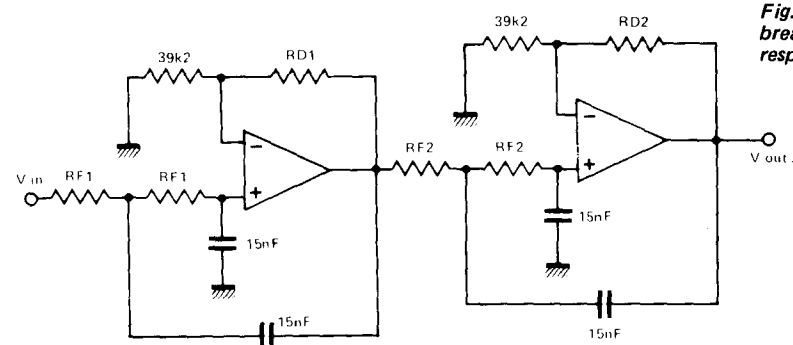|	RF1	RD1	RF2	RD2	GAIN IN dB	COMPONENT TOLERANCE
BESSEL	7k45	3k24	6k60	29k4	5.6	10%
BUTTERWORTH	10k66	5k9	10k66	48k7	8.3	5%
CHEBYSHEV	24k11	42k2	11k20	71k5	15.?	1%

Fig. 8c. Fourth order low-pass filter.

741 as the op amp the frequency limit is about 10 kHz. With an LM 318 the limit can go right out to 200 kHz.

Another problem is the range of values of RF. If RF is too small large currents flow from the op amp and this may affect filter performance. If RF is too large there may be hum pick-up, and dc offset voltages due to bias currents. So keep RF between 1 k and 100 k. If RF appears to need to exceed this range scale the capacitor accordingly.

Active Filter Cookbook

$$f_0 = \frac{1}{2\pi C \sqrt{R1\,R2}}$$

$$Q = \tfrac{1}{2}\sqrt{R1/R2}$$

$$\text{GAIN} = -2Q^2$$

Q	R1	R2	GAIN IN dB
1	5k33	21k32	6 dB
2	2k66	42k66	18.1dB
3	1k77	60k40	25.1 dB
4	1k33	85k33	30.1 dB
5	1k06	106k66	34.0 dB

Fig. 10. A multiple feedback bandpass filter. The centre circuit is normalised for 1 kHz. The table is the design table for this circuit. To change the design frequency change R1 and R2 by an equal factor.

Band-pass filters

Several second order band-pass filters can be cascaded to produce a different response shape, which, like those discussed earlier for low and high-pass filters, can be optimised to give maximum roll-off or maximum pass-band flatness. Such filters do however tend to be difficult to design and so only second order filters will be discussed.

Figure 10 shows a simple band-pass filter known as a multiple feedback circuit. This circuit can provide only low orders of Q (up to about 5). It will probably oscillate if designed to provide higher Q. Note that a high Q implies a large gain at centre frequency. Care must therefore be taken to ensure that the op amp has sufficient band-width to cope.

The design chart is shown in Fig 10 — once again all values are shown for 1 kHz. The design procedure is to first

choose the Q factor and then perform the frequency scaling. For instance if the centre frequency is to be 250 Hz then multiply both R1 and R2 by four.

If a high Q is required then you must use a multiple op amp circuit such as the 'state variable' or 'Bi-Quad' circuits shown later. Both circuits can achieve Qs as high as 500.

State-variable filter

A state-variable filter is shown in Fig 11. This filter has three main features —
1. It can provide a stable high Q
2. It is readily tuned
3. It is versatile, providing band-pass, low-pass and high-pass outputs all at the same time.

The Q of this filter is determined by the ratio of resistors RA and RB where RA/RB=3Q-1. The resonant frequency fc is $1/2\pi RC$

Note that there are two Cs and two RFs in the circuit. If the filter is to be tunable then both RFs should be changed by an equal amount (the RFs can of 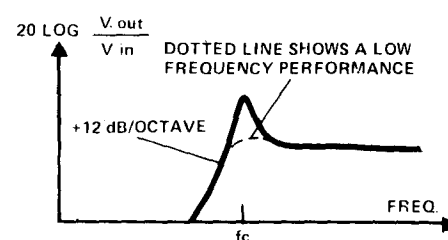course be a dual potentiometer). Note too that Q and fc are independent of each other so as the resonant frequency is changed Q remains constant — and vice versa.

The requirements placed on the op amps in the state-variable filter are less than those for the multiple feedback circuit. The op amps need only have an open loop gain of 3Q at the resonant frequency. Suppose we have a Q of 100 and an fc of 100 kHz. Then the open loop gain is 300, the frequency is 10 kHz and so the gain band-width product needed is 3 MHz.

When using a high Q, care must be taken with signal levels. The gain of the filter is +Q at resonance so if you

are filtering a 1V signal with a Q of 100 you could expect a 100 V output signal!

National Semiconductor manufactures an active filter IC which is a four op amp network which can be used to make up state-variable filters with Qs of up to 500 and frequencies to 10 kHz. The device is designated AF 100.
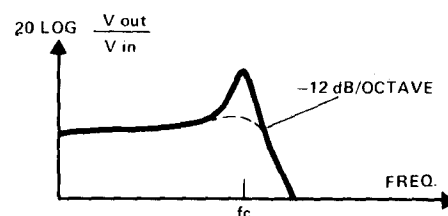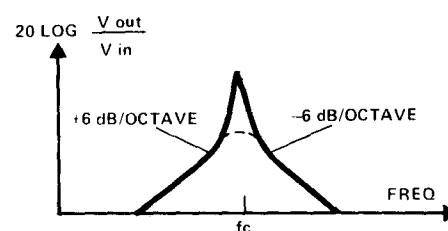

Fig. 11b. The state variable filter is called a universal filter because it can give band-, low- and high-pass outputs — as shown above. Note that all these responses are second order in nature.

Fig. 11a. The state variable filter.

Fig. 12. A Bi-Quad active filter design.

RG

10k66

C
15nF

RF
10k66

15nF

10k   10k

Vin   10k66

BP

LP

$$fc = \frac{1}{2 \cdot CR \cdot F}$$

$$Q = \frac{RQ}{RF}$$

NOMALIZED FOR 1kHz OPERATION
TO FREQUENCY CHANGE IT, CHANGE
ALL THREE RFs BY AN EQUAL FACTOR

Q IS LINEARLY PROPORTIONAL TO fc

$$BANDWITH\ fBW = \frac{fc}{Q}$$

THEREFORE BANDWITH REMAINS
CONSTANT

Bi-Quad filter

Figure 12 shows a Bi-Quad active filter. It looks very similar to the state-variable filter but the minor changes cause it to behave quite differently. Unlike the state-variable filter it has only a band-pass and a high-pass output. The resonant frequency is

$$fc = \frac{1}{2\pi CRF}$$

FREQUENCY RANGE 20Hz 20kHz 18k

741

VOLTAGE CONTROL INPUT
1V/DECADE

PNP DUAL
TRANSISTOR

+12V

120k

10k LIN
COARSE
FREQUENCY POT

10nF

1M   1k

~ 12V
+12V

390k

10k LIN
FINE FREQUENCY POT

POWER SUPPLY
12V
REGULATED

12V

20%,
22k

2%,
22k

−12V

CURRENT CONTROL GENERATOR
CONVERTING LINEAR INPUT
VOLTAGES TO LOG CURRENTS

100k

741   39k

47k

CA 3080

CA 3080

CA3140

LP*

100k   1.F

Vin

100k
INPUT
LEVEL

120R

5V6 5V6   680k

150pF

47k

BP*   120R

120R

150pF

100pF
120R

PRESET Q FOR
MAXIMUM NOTCH
DEPTH

SET OUT OF BAND
REJECTION

33k   10k   33k

TO LP

4M7
10k

27k   SW1a

10k LOG
Q POT

33k

10k

10k   100k   TO BP

SWITCH
SELECTS
SEE NOTE

SW1b

741

NOTCH

741

HP

27k   8k2   47k   47k

A BP,LP,HP FILTER OPERATION
B QUADRATURE OSCILLATOR.
   USE THE BP AND LP OUTPUTS
C NOTCH MODE. USE NOTCH OUTPUT

*DO NOT LOAD THE BP OR
THE LP WITH LESS THAN
10k IMPEDANCE

Fig. 13. The state variable filter can also be made to oscillate (as above). It has a variable resonant frequency, and so becomes a variable frequency oscillator. This circuit produces two low-distortion sinusoids in phase-quadrature: i.e. sine and cosine waveforms.

Active Filter Cookbook

Notch filters

Two notch filters were shown earlier in this article. These worked by mixing a band-pass signal with the original or by mixing the low and high-pass outputs. There are however many other methods of producing a notch response.

The twin-T circuit shown in Fig 14 is very interesting — a notch response is obtained using resistors and capacitors only. However as this is purely a passive device only a low Q can be obtained. The circuit is not used a great deal — perhaps because no less than six components determine its notch frequency. However it is of interest to note that when the twin-T is placed in the feedback loop of a high-gain amplifier a band-pass response will be obtained. If R is made variable it is possible to move the centre frequency, although in doing so the Q varies. This effect has been exploited as the basis of many wah-wah effects units!

All-pass filter

Another way of obtaining a notch is the all-pass filter shown in Fig 16. The frequency response might at first sight make this the most useless of all possible filters — it's flat! However the circuit produces a phase shift which goes from 180°, through 90° at fc — to

Fig. 14. Twin T band notch filter configuration.

$$fc = \frac{1}{2\pi RC} \text{ Hz}$$

WHERE fc = 1kHz
C = 15nF
C2 = 30nF
R = 10k66
R2 = 5k33
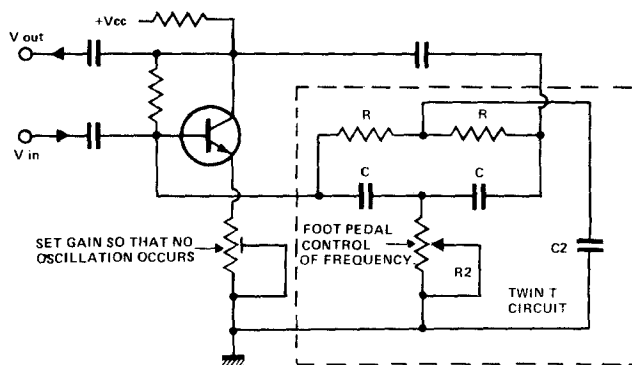
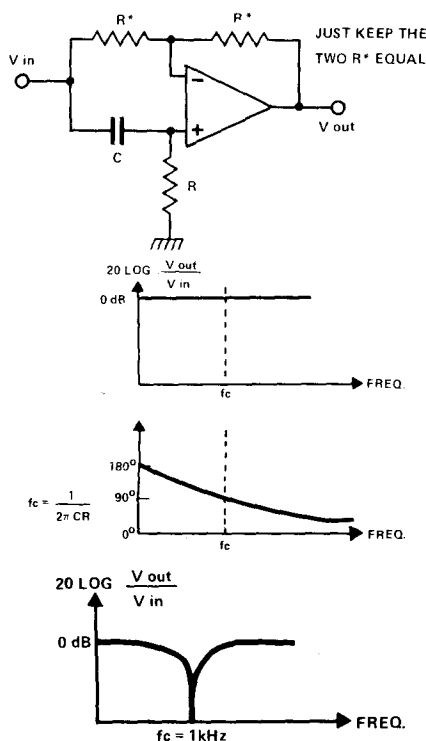Fig. 15. Block diagram of a typical Waa-Waa pedal.

Fig. 16. All-pass filter. At the top is the circuit for such a device. Its frequency and phase responses are shown below it, with the obtainable notch at the bottom.
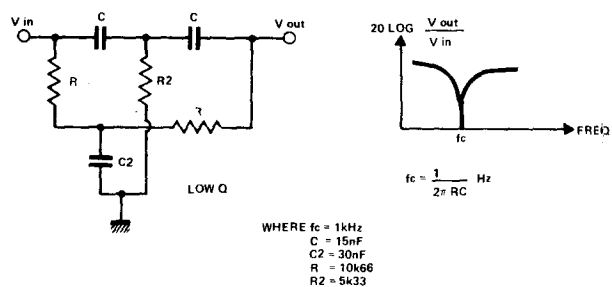
0°. By cascading two such filters the phase shift is doubled: if we then mix the phase-delayed signal with the original a notch response is obtained. This is because at fc the two signals have the same magnitude but opposite phase and so cancel out.

If the notch is to be tunable the RC time constants must be variable. Just one R may be varied if a small range is required — otherwise vary both Rs.

Comb filter

Several notches can be produced by cascading notch filters. This type of filter is called a comb filter. Every notch in the comb requires two all-pass filters. The notches can be made to move up and down in frequency by making the Rs variable. This method is used to produce the 'phasing' effect used by rock groups.

Figure 17 shows a small section of such a unit. Here a CMOS chip is used to provide a matched set of six MOSFETS. A common voltage controls the MOSFET's channel resistance. Thus as the control voltage varies so do the six MOSFET resistors — and the three notches move along the frequency axis in unison.

Another type of comb filter is shown in Fig 18. Here a time delay line is used instead of a phase delay line. This produces a large number of notches linearly spread along the frequency axis, their spacing being determined by the
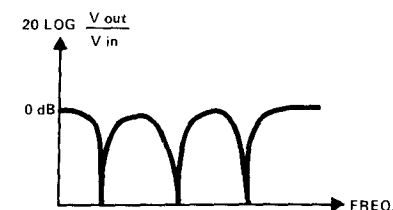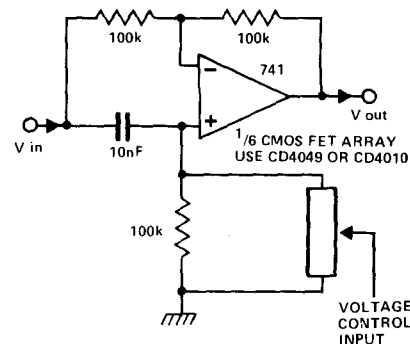
Fig. 17. One section of a comb filter. The response produced by the full (six times above) circuit is shown below the circuit.

delay times.

A bucket-brigade device may be used to implement the time delays (which can of course be made variable). This type of filter is known as a flanger and is used to generate high quality phasing effects. An even more impressive sound can be produced by adding feedback around the delay line: this produces a multi-peak, high-Q filter which makes very interesting 'musical' sounds when swept through its range.
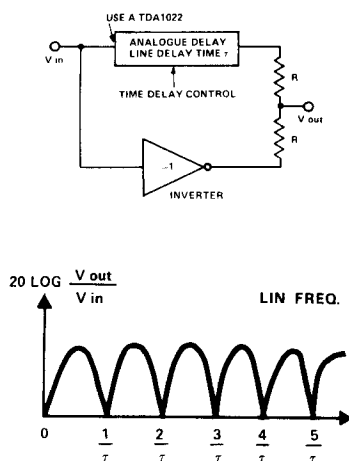
Fig. 18. Alternative method of producing a comb filter using a Mullard delay line.

Fig. 19. Another method of varying the notch frequency, mark/space ratio modulation, and has the advantage of possessing a wide range.

Fig. 20. Simple tone control circuit, with the lift and cut responses shown beneath it.

Variable tuning

A common requirement is for a variable centre or cut-off frequency. This causes problems with filters of orders more than two simply because it is difficult to obtain potentiometers with more than two ganged sections. Lots of companies make them but just try to buy one when it's needed!

One solution is to use mark/space ratio modulation (Fig 19). This has the advantage of offering continuously variable control over a range of about 100:1. Lots of sections may be used and they'll all track. As an example an eighth order four transmission-gates/pack variable frequency filter can be made using a couple of CD4016s. Note though that —

1. The switching waveform must be several times higher in frequency than the highest frequency to be filtered.
2. More circuitry, to generate the switching waveform is required.
3. Switching noise tends to be generated.

Audio circuits

Active filters are widely used in equalising audio signals in applications varying from simple tone controls in hi-fi systems to parametric equalisers in recording studios. Figure 20 shows a commonly used tone control with bass and treble functions. Cut and lift ranges are 20 dB. More flexible control can be obtained by using a multi-band graphic equaliser such as that described as a constructional project in ETI's June 1977 issue.

Testing

Once the process of designing active filters has been reduced to a simple procedure, testing should be made as simple as possible too. The most basic method is to use a swept sinewave oscillator (Fig 21).
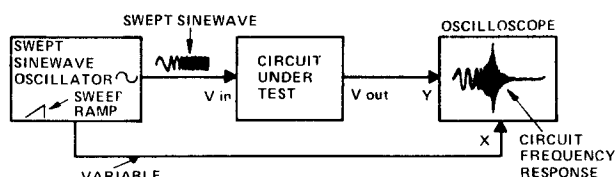
Fig.21. Testing active filters using a sweep oscillation.

An XY oscilloscope is used to display frequencies logarithmically against amplitude (displayed linearly). The ideal display would be log amplitude but this is hard to do. The beauty of this method of testing is that the display is in real time so any changes appear instantly on the oscilloscope. If high Qs or rapid roll-offs at low frequencies are involved, then the sweep time will have to be reduced, otherwise the effects of ringing will 'time smear' the display. The harmonic distortion of the sinewave can be quite large (0.5 to 2.0%) without causing too much of a display problem for most filter designs.

TIMER

APPLICATIONS

DESCRIBED BY R. M. MARSTON
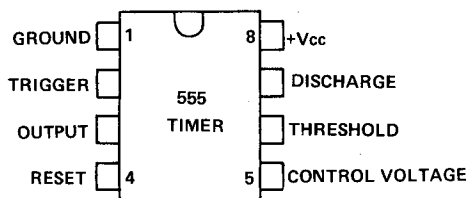
Fig. 1. Outline and pin notations of the standard 8-pin DIL version of the 555 timer I.C.
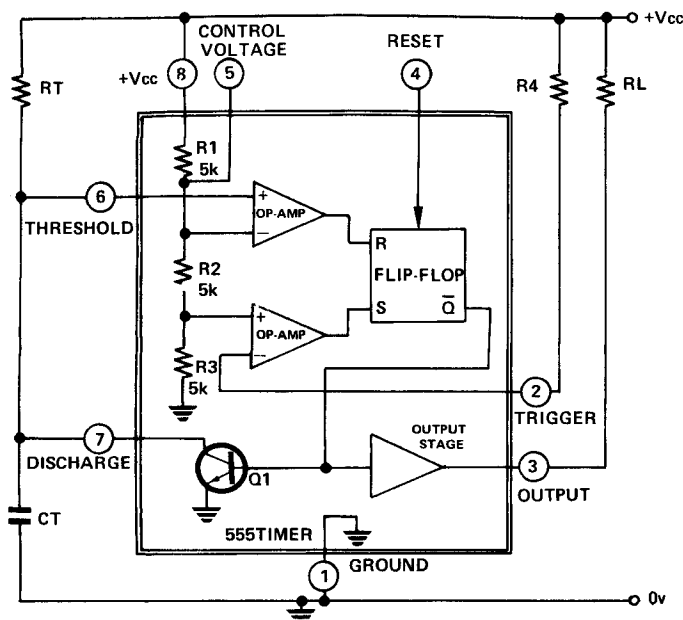
Fig. 2. Functional block diagram (within the square) of the 555 timer i.c., together with the connections for using the i.c. as a basic monostable generator or timer.

THE 555 TIMER is a highly versatile low-cost IC specifically designed for precision timing applications. It can also be used in monostable multi-vibrator, astable multivibrator, and Schmitt trigger applications. The device was originally introduced by Signetics, but is now available from many other manufacturers.

The 555 has many attractive features. It can operate from 4.5v to 16v. Its output can source (supply) or sink (absorb) any load current up to a maximum of 200 mA, and so can directly drive loads such as relays, LED's, low-power lamps, and high impedance speakers. When used in the 'timing' mode, the IC can readily produce accurate timing periods variable from a few microseconds to several hundred seconds via a single R-C network. Timing periods are virtually independent of supply rail voltage, have a temperature coefficient of only .005% per°C, can be started via a TRIGGER command signal, and can be aborted by a RESET command signal.

When used in the monostable mode, the IC produces output pulses with typical rise and fall times of a mere 100 nS. It can be made to produce pulse-width modulated (PWM) pulses in this mode by feeding fixed frequency clock pulses to the TRIGGER terminal and, by feeding the modulation signal to the CONTROL VOLTAGE terminal.

When used in the astable mode both the frequency and the duty cycle of the waveform can be accurately controlled with two external resistors and one capacitor. The output signals can be subjected to frequency sweep control, frequency modulation (FM), or pulse-position modulation (PPM) by applying suitable modulation signals to the CONTROL VOLTAGE terminal of the IC.

HOW IT WORKS

The 555 is available under a variety of specific type numbers but is generally referred to simply as a '555 timer.' The device is available in a number of packaging styles, including 8 and 14-pin dual-in-line (DIL) and 8-pin TO-99 types. Throughout this article all circuits are designed around the standard 8-pin DIL versions of the device.

Figure 1 shows the outline and pin notations of the standard 8-pin DIL version of the 555, and Fig 2 shows the functional block diagram of the same device (within the double lines), together with the connections for using it as a basic monostable generator. The following explanation of device operat-
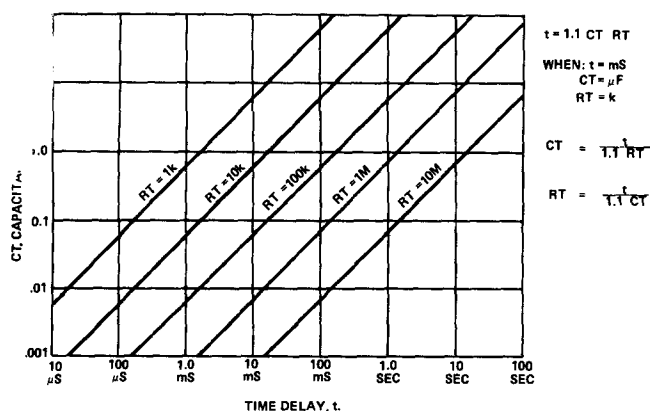
Fig. 3. 555 time delays for different values of resistance and capacitance.

$$t = 1.1 \ CT \ RT$$

WHEN: t = mS
CT = μF
RT = k

$$CT = \frac{t}{1.1 \ RT}$$

$$RT = \frac{t}{1.1 \ CT}$$

ion assumes that the 555 is used in the monostable configuration shown in Figure 2.

The 555 houses two diodes, 15 resistors, and 23 transistors. These components are arranged one voltage-reference potential divider, two voltage, comparator op-amps, one R-S flip-flop, a low-power complementary output stage, and a slave transistor. The voltage-reference potential divider comprises three 5 kΩ resistors in series, and is connected across the supply lines. Consequently, 2/3 Vcc appears at the junction of the upper two resistors of the potential divider, and is fed to one input terminal of the upper voltage-comparator op-amp and 1/3 Vcc appears at the junction of the two lower resistors of the potential divider, and is fed to one input terminal of the lower voltage-comparator op-amp. The outputs of the two comparators control the R-S flip-flop, which in turn controls the states of the complementary output stage and the slave transistor. The state of the flip-flop can also be influenced by signals applied to the pin 4 RESET terminal.

When the monostable or timing circuit of Fig 2 is in its quiescent state, the pin 2 TRIGGER terminal of the chip is held high via R1. Q1 is driven to saturation and forms a short circuit across external timing capacitor CT, and the pin 3 output terminal of the IC is driven to the low state. The monostable action can be unitiated by applying a negative-going trigger pulse to pin 2. As this pulse falls below the 1/3 Vcc, reference value of the built-in potential divider the output of the lower voltage comparator op-amp changes state and causes the R-S flip-flop to switch over. As the flip-flop switches over it cuts off Q1 and drives the pin 3 output of the chip to the high state.

As Q1 cuts off it removes the short from timing capacitor CT, so CT starts to charge exponentially towards the supply

rail voltage until eventually the voltage across CT reaches 2/3 Vcc. At this point the upper voltage comparator op-amp changes state and switches the R-S flip-flop back ot its original condition, so Q1 turns on, rapidly discharging CT, and simultaneously the pin 3 output of the IC reverts to its low state. The monostable operating sequence is then complete. Note that, once triggered, the circuit cannot respond to additional triggering until the timing sequence is complete, but that the sequence can be aborted at any time by feeding a nagative-going pulse to pin 4.

The delay time of the circuit, in which the pin 3 output is high, is given as

$$t = 1.1 \ R \quad C$$

where t = mS, RT = kΩ, and CT = μF. Figure 3 shows how delays from 10μS to 100 seconds can be obtained by selecting suitable values of CT abd RT in the range .001μF to 100 μF and 1 kΩ to 10 MΩ or greater than 20 MΩ, and capacitor CT must always be a low-leakage component. Note that the timing period of the circuit is virually independent of the supply voltage but that the period can be varied by applying a variable resistance or voltage between the ground and pin 5 CONTROL VOLTAGE terminals of the chip. This facility enables the periods to be externally modulated or compensated.

The pin 3 output terminal of the IC is normally low, but switches high during the active monostable sequence. The output can either source or sink currents up to a maximum of 200 mA, so external loads can be connected between pin 3 and either the positive supply rail or the ground rail, depending on the type of load operation that is required. The output switching rise and fall times are typically about 100 nanoceconds. Having cleared up these points, let's now go on and look at some practical applications of the 555 timer I.C.

50 SECOND TIMER

This 50 second timer or pulse generator gives a direct voltage output at pin 3 which is normally low, but goes high for the duration of the timing period. Optional components R4 and LED (shown dotted) give a visual indication of the timer action. The circuit works in the same basic way as already described, except that the timing action is initiated by momentarily shorting pin 2 to ground via START switch S1 Note from the circuit waveforms that a fixed-period output pulse is available at pin 3 and an exponential sawtooth with an identical period is available at pin 7: The sawtooth waveform has a high output impedance.

The basic timer circuit of Fig 4 can be varied in a number of ways. The timing period can be made variable between approximately 1 . 1 seconds and 110 seconds by replacing
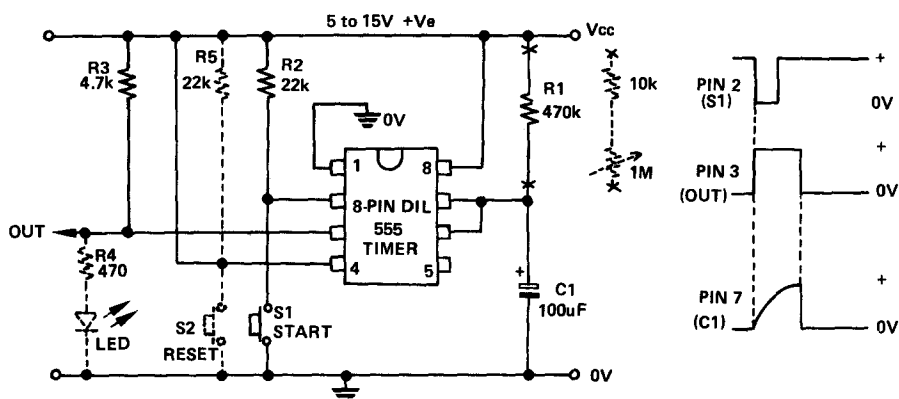

Fig. 4. Circuit and waveforms of simple manually-triggered 50 second timer or pulse generator.
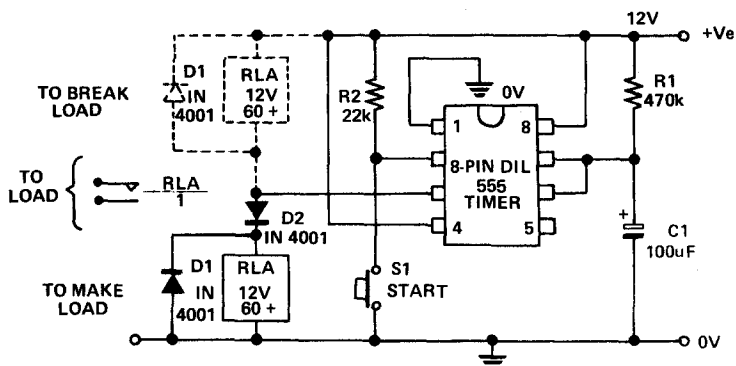
555 TIMER APPLICATIONS

Fig. 5. Relay-output timer makes or breaks connection to load for pre-set period of 50 seconds when S, is momentarily operated.

R1 with a 10 kΩ fixed resistor and a 1 MΩ variable resistor in series.

The period can be further varied, if required, by switch-selecting decade values of timing capacitance. The dotted section shows how the circuit can be provided with a RESET facility, so that a timing period can be aborted at any time, by taking pin 4 to the positive supply rail via resistor R5 and wiring RESET switch S2 between pin 4 and ground.

The timing circuit of Fig 4 can be used to drive non-inductive loads at currents up to 200 mA directly. They can be used to drive inductive relay loads by using the basic connections shown in Fig 5.

The (Fig 5) circuit is designed to apply a connection to a normally-off external load for a pre-set period of 50 seconds when START switch S1 is momentarily closed. The relay is normally off, but turns on for the 50 second period when the timing cycle is initiated. D2 is wired in series with the relay

coil to counteract the slight residual voltage that ap at pin 3 of the IC under the OFF condition and thus ens that the relay turns fully off. The dotted section shows ho this circuit can be used to switch off a normally-on load.

Note in Fig 5 and all other relay-output circuits describeo here, that the relays used can be any 12 volt types that draw ON currents of less than 200 mA, e.g., that have coil resistances greater than 60 Ω.

The basic relay-driving timer circuit of Fig 5 can be adapted for use in a variety of useful applications. Some typical examples are shown in Figs 6 to 9.

Figure 6 shows the practical circuit of a relay-output general-purpose timer that covers 0.9 seconds to 100 seconds in two decade ranges. The circuit has a RESET facility provided via S2 , so that timing periods can be aborted part way through a cycle if necessary. A noteworthy feature of this circuit is that the maximum timing periods of each decade range of the timer can be precisely pre-set via R5 or R6 , which effectively shunt the built-in potential divider of the 555 and thus influence the timing periods: This facility enables the circuit to give precise timing periods even when wide-tolerance timing capacitors are used.

To set up the Fig 6 circuit, first set R1 to maximum value, set RANGE switch S3 to position 1, activate START switch S1, and adjust R5 to give a timing period of precisely 10 seconds. Next, set S3 to position 2, activate START switch S1, and adjust R6 to give a timing period of precisely 100 seconds. All adjustments are then complete, and the timer is ready for use.

DELAYED HEADLIGHT TURN-OFF

Figure 7 shows the practical circuit of an automatic delayed-turn-off headlight control system for auto-mobiles. This facility enables the owner to use the car lights to illuminate his path for a pre-set time after parking as he leaves the garage or walks along a driveway, etc. The circuit does not interfere with normal headlight operation under actual driving conditions. It works as follows.

When the ignition switch is turned to the ON position current is fed to the coil of the relay via D3 and the 12 volt supply rail, so the relay turns on and contacts RLA/1 close. As the contacts close they connect the 12 volt supply rail, so the relay turns on and the headlight switch. Thus, under this 'ignition on' condition the headlights operate in the normal way. Note that, since one side of C2 is connected
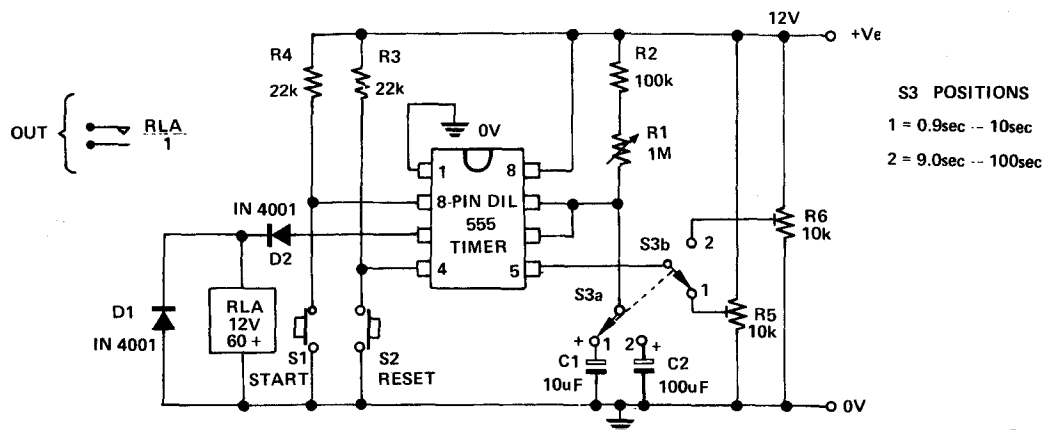

Fig. 6. Relay-output general-purpose timer covers 0.9 sec to 100 sec in two decade

Fig. 7. Automatic delayed-turn-out headlight control system for automobiles.

directly to the positive supply rail and the other side is taken to the positive rail via R2, the capacitor is fully discharged under this condition.

The moment that the iginition switch is turned to the OFF position the D3-derived current supply to the relay coil is broken, and simultaneously a negative-going trigger pulse is fed to pin 2 of the 555 as the C2-R3 junction drops to ground volts and C2 charges up. Relays are inherently slow-acting devices, so contacts RLA/1 do not open instantaneously as the ignition switch is turned off. Conversely, the 555 is a very fast triggering device, and the instant that the trigger pulse is generated, via the turn-off action of the ignition switch a timing cycle is initiated and current is fed to the relay coil via output pin 3 of the IC as it goes high. Thus the relay remains on for a pre-set period after the ignition connected to the headlight switch for the duration of this period. With the component values shown this period is roughly 50 seconds.

At the end of the 50 second timing period, pin, 3 of the 555 switches to the low state and the relay turns off. As it does so, contacts RLA/1 open and remove the supply from the timer and the headlight switch, and the headlights turn off. The operating sequence is then complete.

Readers may care to note that the above system of operation is consistent with the practice adopted in many modern vehicles of feeding the headlight switch via the ignition switch, so that the headlights operate only when the ignition is turned on. On older types of vehicle, where headlight operation is

independent of the ignition switch, a manually-triggered delayed-turn-off headlight or spotlight control facility can be obtained by using the circuit shown in Fig 8. The action of this circuit is such that, if the vehicle is parked with its lights off, they turn on for a pre-set 50 second momentarily closed, and at the end of this period turn off again automatically.

The (Fig 8) circuit uses a relay with two sets of normally-open relay contacts. The timing sequence is initiated by momentarily closing push-button switch S1. Normally, both S1 and the relay contacts are open, so zero power is fed to the timer circuit and the lights are off. C2 discharged under this condition.

When S1 is momentarily closed power is fed directly to the relay coil, and the relay turns on. As the relay turns on, contacts RLA/2 close and apply power to the vehicle lights and contacts RLA/1 close and apply power to the timer circuit, but pin 2 of the IC is briefly tied to ground via C2 and R3 at this moment, so a negative trigger pulse is immediately fed to pin 2 and a timing cycle is initiated. Consequently, pin 3 of the 555 switches high at the moment that the relay contacts close, and thus locks the relay into the ON condition irrespective of the subsequent state of START switch S1 so the lights remain on for the duration of the 50 second timing cycle. At the end of the timing cycle pin 3 of the IC switches to the low state, so the relay turns off and contacts RLA/1 and RLA/2 open, disconnecting power from the timing circuit and the lights. The operating sequence is then complete.
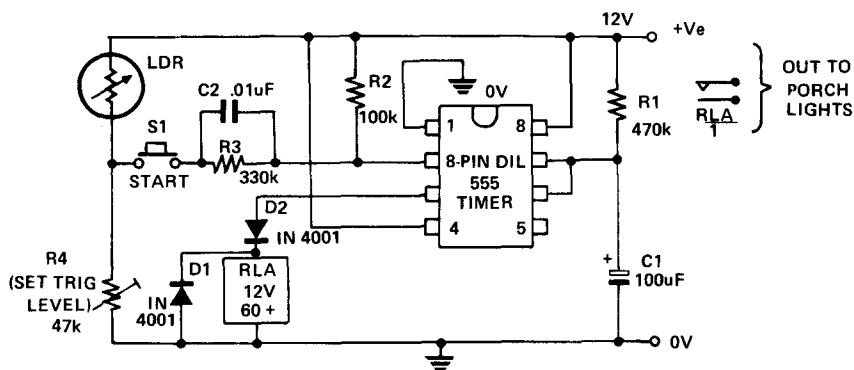

Fig. 8. Manually-triggered delayed-turn-off head- or spot-light control system for automobiles.
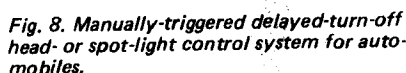
Fig. 9. Automatic porch light turns on for a pre-set period only when triggered at night.

PORCH LIGHT

Finally, to conclude this 'Timer Circuits' section of the 555 story, Fig 9 shows the circuit of a relay-output automatic porch light control unit that turns the porch lights on for a pre-set 50 second period only when suitably triggered at night-time or under 'dark' conditions: The circuit is triggered via switch S1, which may take the form of a microswitch activated by a porch gate or a pressure-pad switch activated by body weight and concealed under a porch mat or rug.

The operation of the Fig 9 circuit relies on the fact that for correct timer operation the negative-going trigger pulse that is fed to pin 2 of the IC must fall below the internally-controlled '1/3 Vcc' voltage value of the 555. If the trigger pulse does not fall below this value, timing cycles can not be initiated by the trigger signal.

In this design, light-dependent resistor LDR and preset resistor R4 are wired in series as a light-dependent potential divider. One side of switch S1 is taken to the output of this postential divider, and the other side of the switch is taken to pin 2 of the IC via the C2-R3 combination. Under bright or daylight conditions the LDR acts as a low resistance, so a high voltage appears at the output of the potential divider. Consequently, the act of closing S1 causes a voltage pulse much higher than '1/3 Vcc' to be fed to pin 2 of the chip, so the timer is not triggered via S1 under the 'daylight' condition.

Conversely, the LDR acts as a high resistance under dark or 'night' conditions, so a low voltage appears at the output of the potential divider. Consequently, the act of closing S1 causes a voltage pulse much lower than '1/3 Vcc' to be fed to pin 2 of the IC, so the time circuit is triggered via S1 under the 'night' condition.

In practice, the LDR can be any cadmium-sulphide photo-cell that presents a resistance in the range 1 kΩ to 100 kΩ under the required minimum 'dark' turn-on condition, and R4 can be adjusted to preset the minimum 'dark' level at which the circuit will trigger. Note that the trigger signal is fed to pin 2 of the IC via the C2-R3 combination, which act as a trigger signal conditioning network that effectively isolates the d.c. component of the LDR-R4 potential divider from the trigger pin of the IC.

MONOSTABLE PULSE GENERATOR CIRCUITS

All the 555 timer circuits that we have looked at so far act essentially as monostable multivibrators or pulse generators. The 555 can be used as a conventional electronically-triggered monostable multivibrator or pulse generator by feeding suitable trigger signals to pin 2 and taking the pulse output signals from pin 3. The IC can be used to generate good output pulses with periods from 5 μs to several hundred seconds. The maximum usable pulse repetition frequency is approximately 100 kHz.

The trigger signal reaching pin 2 must be a carefully shaped negative-going pulse. Its amplitude must switch from an OFF value greater than 2/3Vcc to an ON value less than 1/3 Vcc (triggering actually occurs as pin 2 drops through the 1/3 Vcc value). The pulse must have a width greater than 100 ns but less than that of the desired output pulse, so that the trigger pulse is removed by the time the monostable period terminates.

One way of determining a suitable trigger signal for the 555 monostable circuit is to convert the input signal to a good square wave that switches between ground volts and the full positive supply rail voltage, and then couple this square wave to pin 2 of the IC via a simple short time-constant C-R dif-
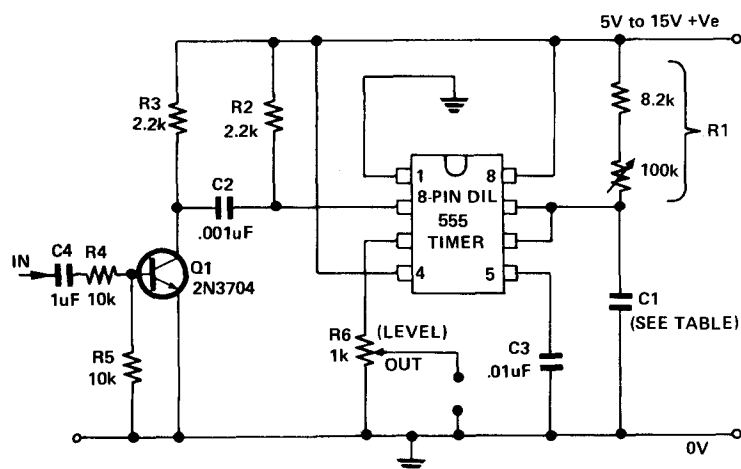
C1 value	Pulse Width Range
10μF	90 ms – 1.2s
1μF	9 ms – 120ms
0.1μF	0.9 ms – 12 ms
0.01μF	90μs – 1.2 ms
0.001μF	9μs – 120 μs

Fig. 10a: Simple add-on pulse generator is triggered by rectangular input signals: circuit can be used at trigger frequencies up to 100 kHz.
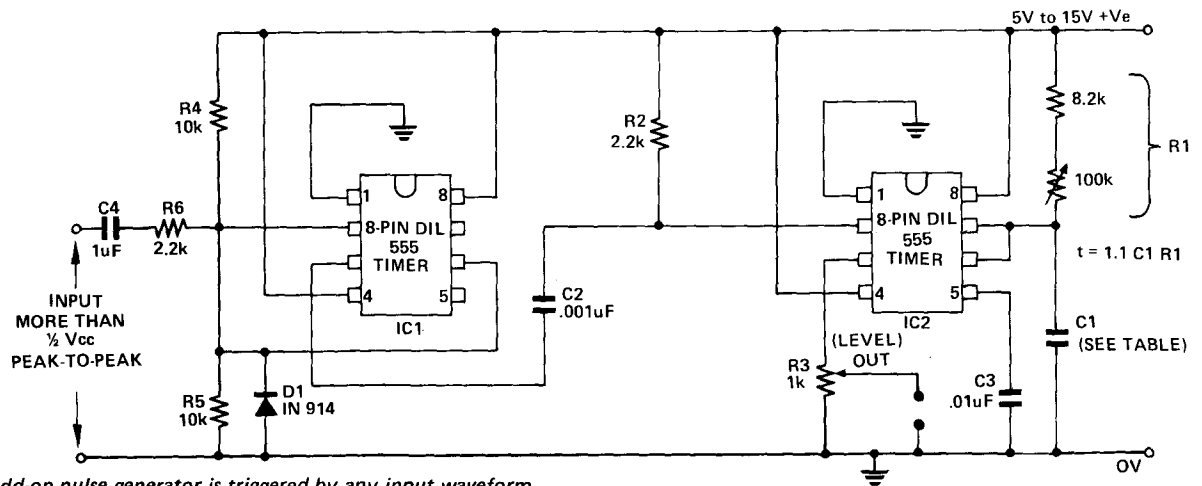
Fig. 10b: Improved add-on pulse generator is triggered by any input waveform.

ferentiating network, which converts the leading or trailing edges of the square wave into suitable trigger pulses. Figure 10a shows a practical circuit that uses this basic principle, but is intended for use only with input signals that are already of square or pulse form.

Here, transistor Q1 converts the rectangular input signal into a signal that switches between the ground and positive voltage rails, and the resulting signal is fed to pin 2 via the differentiating network. The circuit can be used as an add-on pulse generator in conjunction with an existing square or pulse generator. Variable-amplitude output pulses are available from pin 3 via variable potential divider R6. The output pulse widths can be varied over more than a decade range via R1, and can be switched in overlapping decade ranges by using the values of C1 listed in the table. With the component values shown the pulse width is fully variable from 9μs to 1.2 seconds. Note that C3 is used to decouple the pin 5 CONTROL VOLTAGE terminal and improve the circuit stability.

Figure 10b shows how the above circuit can be modified so that it can be driven from any type of input waveform, including sine waves. Here, IC1 is connected as a simple Schmitt trigger, which converts all input signals into rectangular output signals, and these rectangular signals are used to

drive the IC2 monostable circuit in the same way as described above. The Fig 10b circuit can thus be used as an add-on pulse generator in conjunction with an existing waveform generator of any type that produces output signals with peak-to-peak amplitudes greater than 1/2 Vcc.

Figure 11 shows two basic monostable pulse generators connected in series to make a delayed pulse generator. IC1 is used as a Schmitt trigger and IC2 controls the delay width and IC3 determines the output pulse width: The final output pulse appears some delayed time after the initial application of the trigger signal. This circuit can be made into a self-contained instrument by building it into the same cabinet as a simple square wave generator which can be used to provide the necessary drive signals.

Any number of basic monostable pulse generators can be wired in series to give a sequential form of operation. Figure 12 for example, shows the circuit and wave-forms of a 3-stage sequential generator, which can be used to operate lamps or relays, etc., in a pre-programmed time sequence once an initial START command is given via push-button switch S1. Note that the pin 4 RESET terminal of all ICs are shorted together and positively biased via R7, and that these terminals can be shorted to ground via SET switch S2: This SET switch
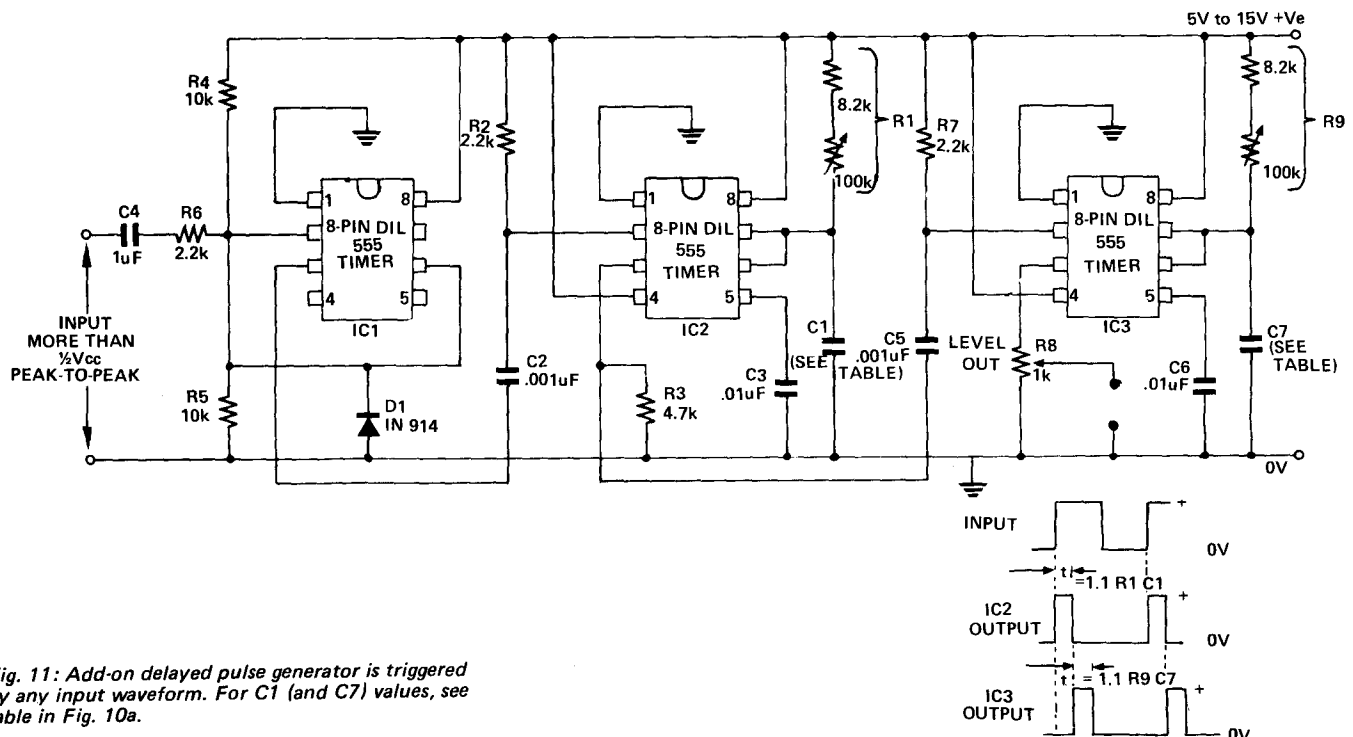

Fig. 11: Add-on delayed pulse generator is triggered by any input waveform. For C1 (and C7) values, see table in Fig. 10a.
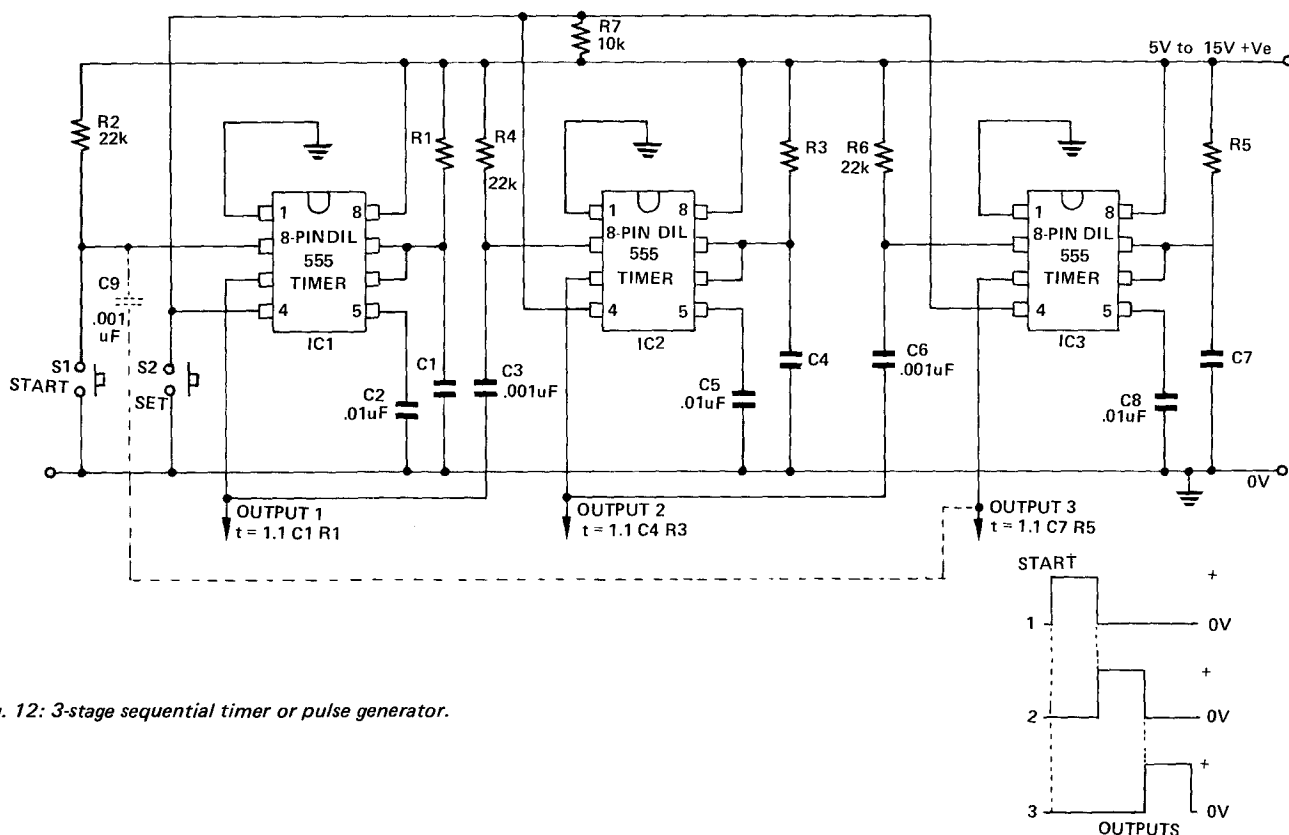
555 TIMER APPLICATIONS

Fig. 12: 3-stage sequential timer or pulse generator.

should be closed at the moment that power is first applied to the circuit, to ensure that none of the ICs falsely triggered at this moment.

Finally, three or more monostable circuits can be connected, via C9, in a continuous loop, with the output of the last monostable feeding back to the input of the first monostable, to form a 'chaser' circuit in which the sequential action repeats to infinity. This type of circuit can be used to drive lamp or LED displays, etc. Note that the circuit is again provided with the S2 SET facility, so that the circuit can be emptied at the moment that power is first applied.

ASTABLE MULTIVIBRATOR CIRCUITS

Figure 13 shows a basic 1kHz astable multivibrator, together with the formulas that define the timing of the circuit. Note that TRIGGER pin 2 of the chip is shorted to the pin 6 THRESHOLD terminal, and that timing resistor R2 is wired between pin 6 and DISCHARGE pin 7.

When power is first applied to the circuit C1 starts to charge exponentially (in the normal monostable fashion) via the series R1-R2 combination, until eventually the C1 voltage rises to 2/3 Vcc. At this point the basic monostable action terminates and DISCHARGE pin 7 switches to the low state. C1 then starts to discharge exponentially into pin 7 via R2, until eventually the C1 voltage falls to 1/3 Vcc, and TRIGGER pin 2 is activated. At this point a new monostable timing sequence is initiated, and C1 starts to recharge towards 2/3 Vcc via R1 and R2. The whole sequence then repeats add infinitum, with C1 alternately charging towards 2/3 Vcc via R1-R2 and discharging towards 1/3 Vcc via R2 only.

Note in the above circuit that, if R2 is very large relative to R1, the operating frequency of the circuit is determined essentially by the R2 and C1 values, and that a virtually symmetrical output waveform is generated. The graph of Fig 14 shows the approximate relationship between frequency and the C1-R2 values under the above condition. In practice, the R1-R2 values of the circuit can be varied from 1 kΩ up to tens of megohms. Note, however, that R1 has a significant effect on the total current consumption of the circuit, since pin 7 of the IC is virtually grounded during half of the timing sequence. Also note that the duty cycle or mark/space ratio of the circuit can be pre-set at a non-symmetrical value, if required, by suitable choice of the R1 and R2 values.

The basic circuit of Fig 13 can be usefully modified in a number of ways. Fig 15, for example, shows how it can be made into a variable-frequency square wave generator by replacing R2 with a fixed and variable resistor in series. With the compnenet values shown the frequency can be varied over the approximate range 650 Hz-7.2 kHz via R2.

Figure 16 shows how the circuit can be further modified so that its MARK and SPACE periods are independently variable over the approximate range 7.5μs to 750μs. Here, timing capacitor C1 alternately charges via R1-R2-D1 and discharges via R3-R4-D2.

Figure 17 shows how the circuit can be additionally modified so that it acts as fixed-frequency square wave generator with a mark/space ratio or duty cycle that is fully variable from 1% to 99%. Here, C1 alternately charges via R1 and the top half of R2 and via D1, and discharges via D2-R3 and the lower half of R2. Note that the sum of the two timing periods
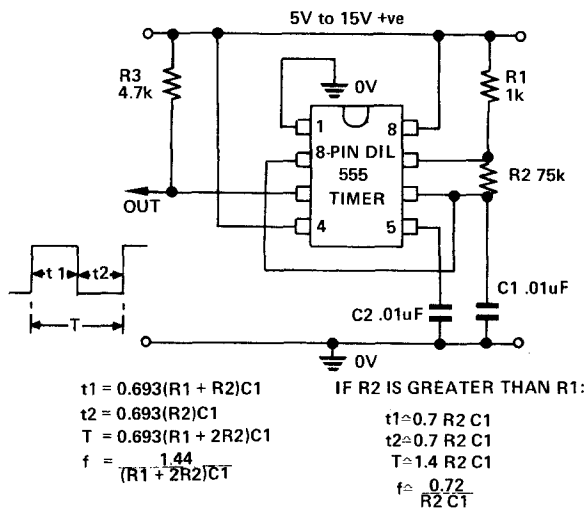
$t1 = 0.693(R1 + R2)C1$
$t2 = 0.693(R2)C1$
$T = 0.693(R1 + 2R2)C1$
$f = \dfrac{1.44}{(R1 + 2R2)C1}$

IF R2 IS GREATER THAN R1:

$t1 \simeq 0.7\,R2\,C1$
$t2 \simeq 0.7\,R2\,C1$
$T \simeq 1.4\,R2\,C1$
$f \simeq \dfrac{0.72}{R2\,C1}$

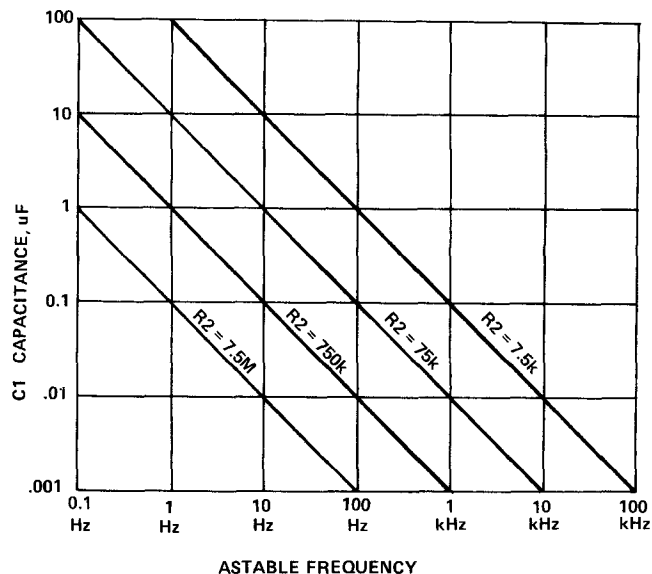Fig. 13: Basic circuit of 1KHz astable multivibrator, with timing formulas.

ASTABLE FREQUENCY

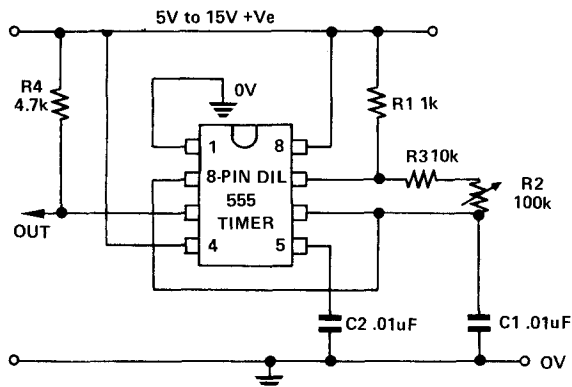Fig. 14: Approximate relationship between C1, R2, and frequency when R2 is large relative to R1.

Fig. 15: Variable frequency square wave generator covers the range 650Hz — 7.2 kHz approximately.
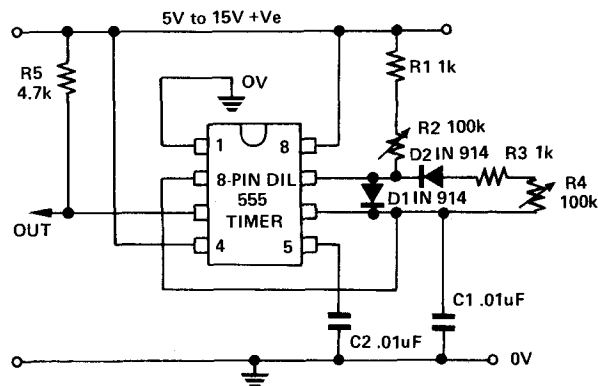
Fig. 16: Astable multi with mark and space periods independently variable over the approximate range 7.5µs to 750µs.

is virtually constant, so the operating frequency is almost independent of the setting of R2.

GATING A 555 ASTABLE

The 555 astable circuit can be gated ON or OFF, via either a switch or an electronic signal, in a variety of ways. Figs 18 and 19 show two basic ways of gating the IC via a switch.

In Fig 18 the circuit is gated via the pin 4 RESET terminal. The characteristic of this terminal is such that, if the terminal is biased significantly above a nominal value of 0.7 volts, the astable is enabled, but if the terminal is biased below 0.7 volts by a current greater than 0.1 mA (by taking the terminal to ground via a resistance less than 7 kΩ, for example) the astable is disabled and its output is grounded. Thus, the Fig 18 circuit is normally on but can be turned off by closing S1 and shorting pin 4 to ground, while the circuit shown in dotted lines is normally gated off via R4 but can be turned on by closing S2 and shorting pin 4 to the positive supply rail. These circuits can alternatively be gated by applying suitable electronic signals directly to pin 4.

The Fig 19a and 19b circuits are gated via the pin 2 TRIGGER and pin 6 THRESHOLD terminals. The characteristic
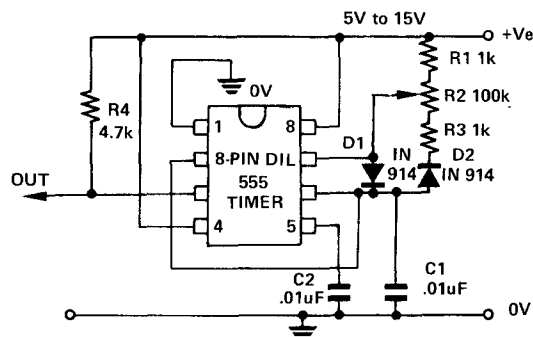

Fig. 17: Astable multi with duty cycle variable from 1 to 99% with frequency approximately constant at 1.2 kHz.
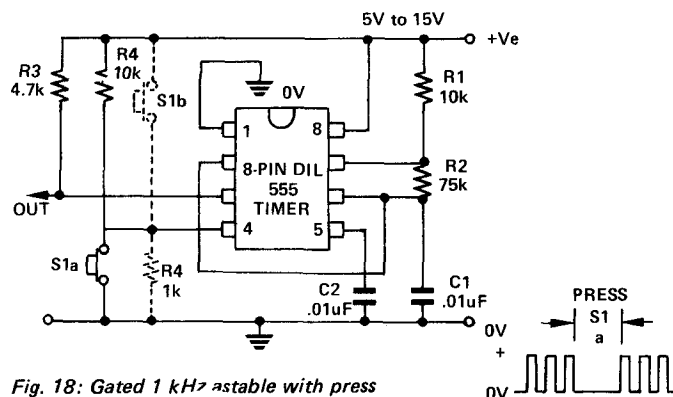
555 TIMER APPLICATIONS

Fig. 18: Gated 1 kHz astable with press

here is such that the circuit functions as a normal astable only as long as pin 6 is free to swing up to 2/3 Vcc and pin 2 is not biased below 1/3 Vcc. If these pins are simultaneously driven below 1/3 Vcc the astable action is immediately terminated and the output is driven to the high state. Thus, the Fig 19a circuit is normally on but turns off when S1 is closed. Note that an electronic signal can be used to gate the circuit by connecting a diode as indicated and eliminating S1. In this

case the circuit will gate off when the input signal voltag reduced below 1/3 Vcc.

The Fig 19b circuit is connected so that it is normall gated off by saturated transistor Q1, but can be gated o by closing S1 and thus turning the transistor off. This cir cuit can be gated electronically by eliminating R5 and S1 and applying a gating signal to the base of Q1 via a 10 kΩ limiting resistor. In this case the astable turns off when the input signal is high, and turns on when the input signal is reduced below 0.7 volts or so.

All the 555 astable circuits that we have looked at can be subjected to frequency modulation (FM) or pulse-position modulation (PPM) by simply feeding a suitable modulation signal to pin 5. This modulation signal can take the form of an ac signal that is fed to pin 5 via a blocking capacitor, as in the case of Fig 20a or a dc signal that is fed directly to pin 5, as in the case of Fig 20b. The action of the chip is such that the voltage on pin 5 influences the width of the 'mark' pulses in each timing cycle, but has no influence on the 'space' pulses. Thus, since the signal on pin 5 influences the position of each 'mark' pulse in each timing cycle, this terminal provides pulse-position modulation (PPM), and, since the signal influences the total period of each cycle (and thus the frequency of the output signal), the terminal also provides frequency modulation (FM). These facilities are useful in special waveform generator applications, as is shown in the next section.
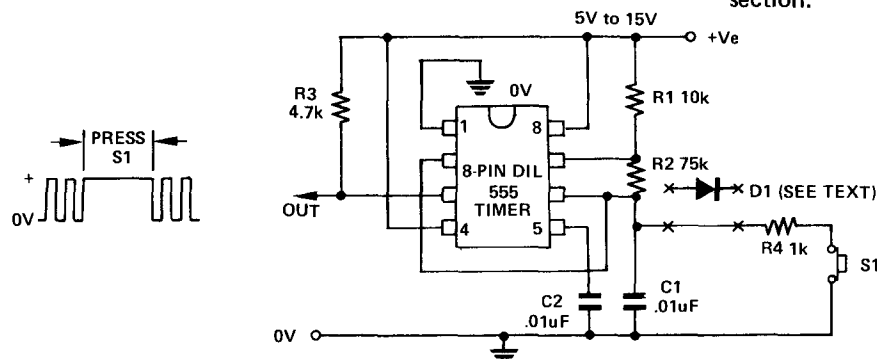

Fig. 19A; Alternative gated 1 kHz astable with 'press to turn-off' operation.
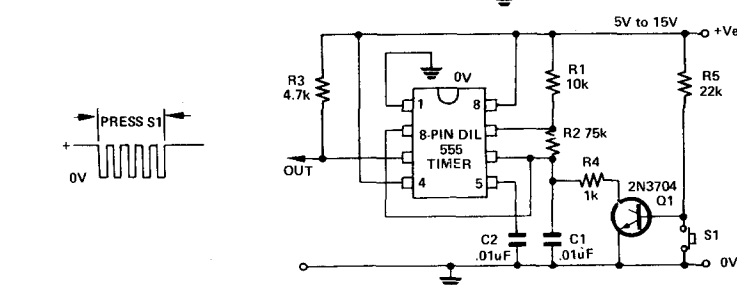

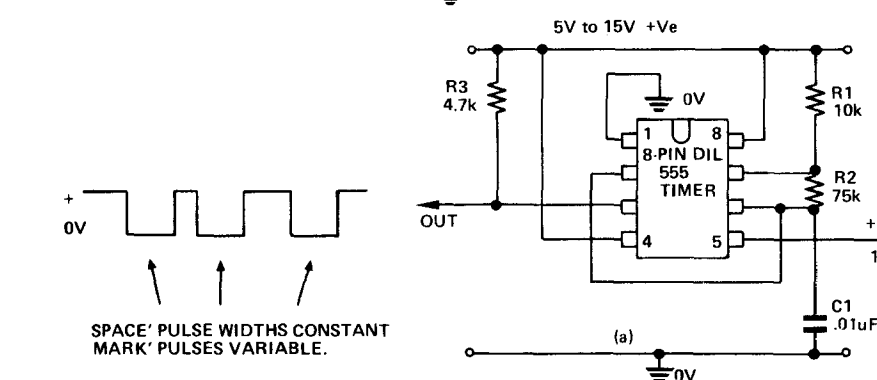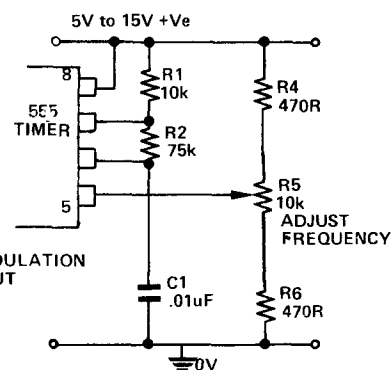Fig. 19B: Alternative gated 1 kHz astable with 'press to turn-off' operation.

Fig. 20: Alternative ways of obtaining frequency or pulse-position modulation (FM or PPM) from the 555 astable circuit.

SPACE' PULSE WIDTHS CONSTANT
MARK' PULSES VARIABLE.

555 TIMER APPLICATIONS

DESCRIBED BY R. M. MARSTON

MISCELLANEOUS ASTABLE APPLICATIONS

The 555 astable multivibrator has three outstanding advantages over other types of astable circuit. First, its frequency can be varied over a wide range via a single resistive control. Second, its output has a low impedance and can source or sink current up to 200 mA. Finally, its operating frequency can readily be modulated by applying a suitable signal to pin 5 of the IC. These features make the device exceptionally versatile, and it can be used in a vast range of practical applications of interest to both the amateur and professional user.

Morse Practice Oscillator

Figure 21 shows how the 555 timer IC can be used as a morse-code practice oscillator. The circuit acts as a normal astable, with frequency variable over the approximate range 300 Hz — 3 kHz via TONE control R3. The 'phone volume is variable via R5, and the 'phones can have any impedance from a few ohms up to megohms. The circuit draws zero quiescent current, since the normally-open morse key is used to connect the circuit to the positive supply rail, which can have any value in the range 5 volts to 15 volts.

Figure 22 shows how the 555 astable circuit can be used in LED flasher applications. This circuit operates at approximately 1 Hz, and has a single LED. The Fig 22 circuit has a single LED output, the dotted section shows how a second may be added, such that one LED is on while the other is off, and vice versa. Any types of LED's can be used in this circuit. Series resistors R1 or R4 determine the ON current of each LED.

Figure 23 shows how the Fig 22 circuit can be modified to

Fig.21. Code-practice oscillator with variable tone (300 - 3kHz) and variable volume.
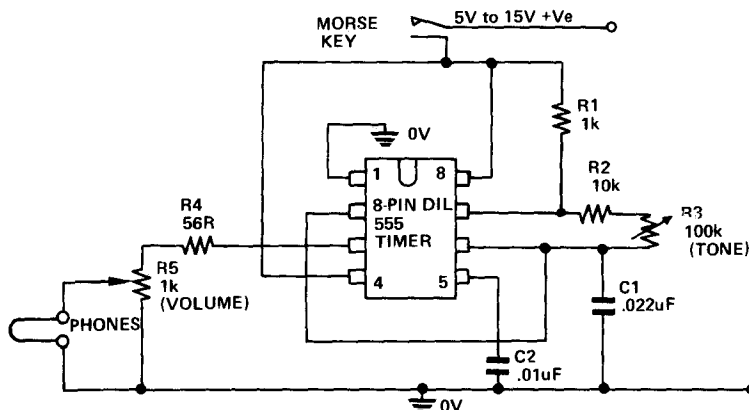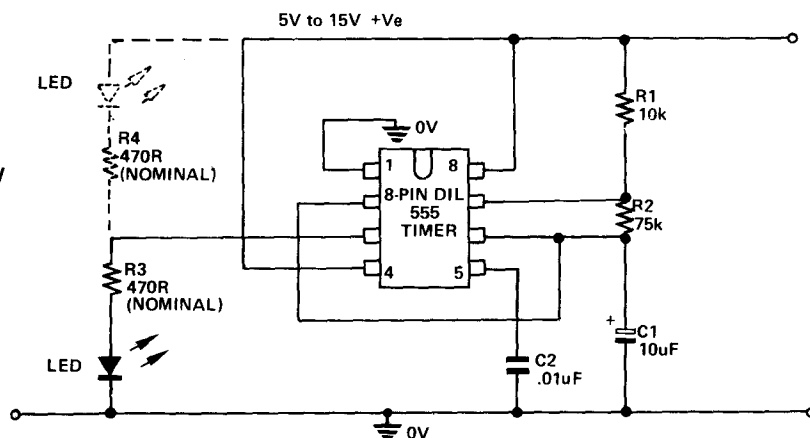

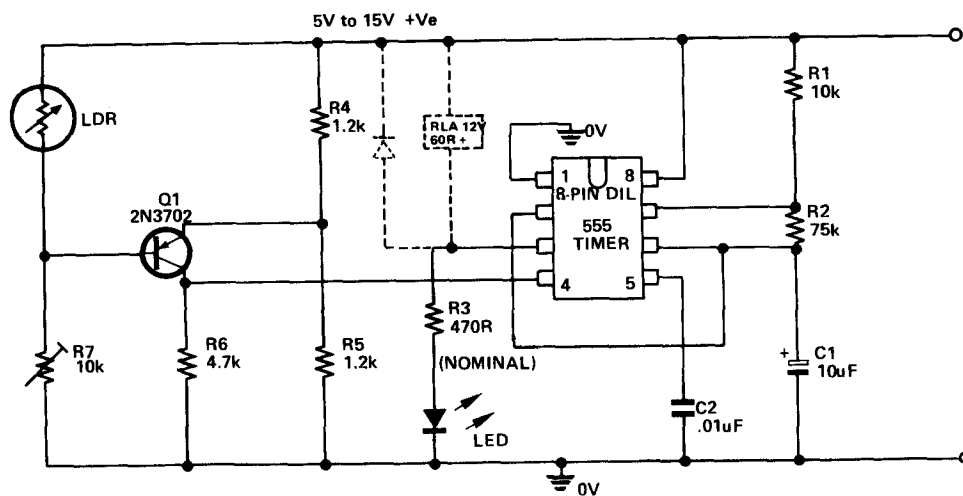Fig.22. Single and dual output LED flashers give symmetrical 1 Hz outputs.

555 TIMER APPLICATIONS

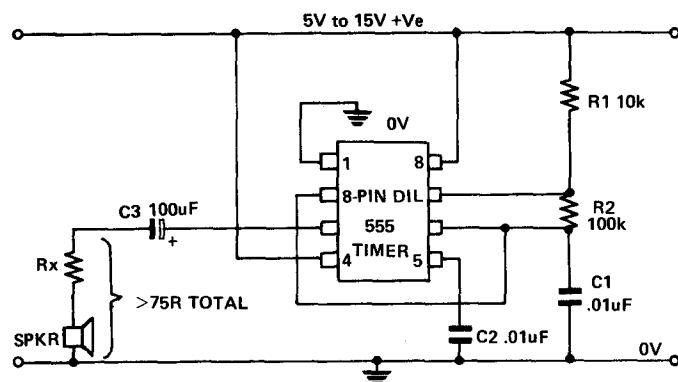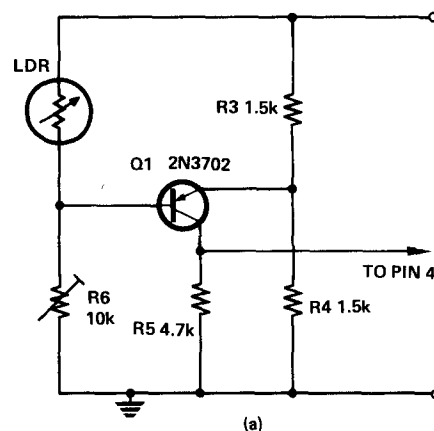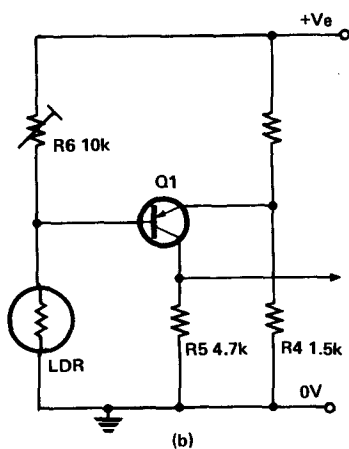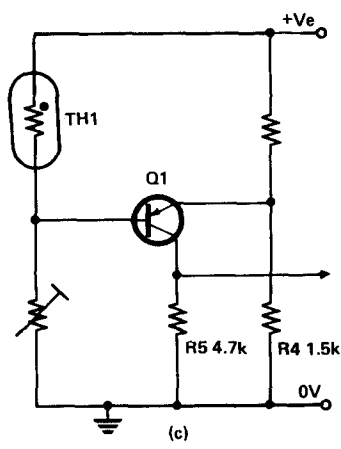Fig.23. Automatic (dark-activated) flasher.

Fig.24. 800Hz monotone alarm call generator
a) activated by dark, b) by light, c) by under-
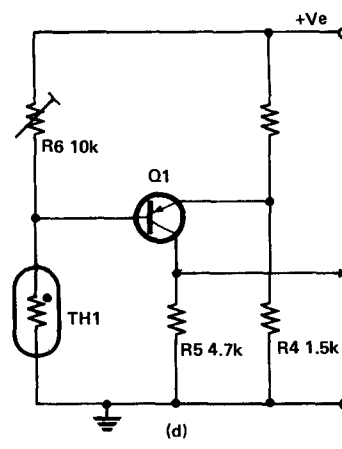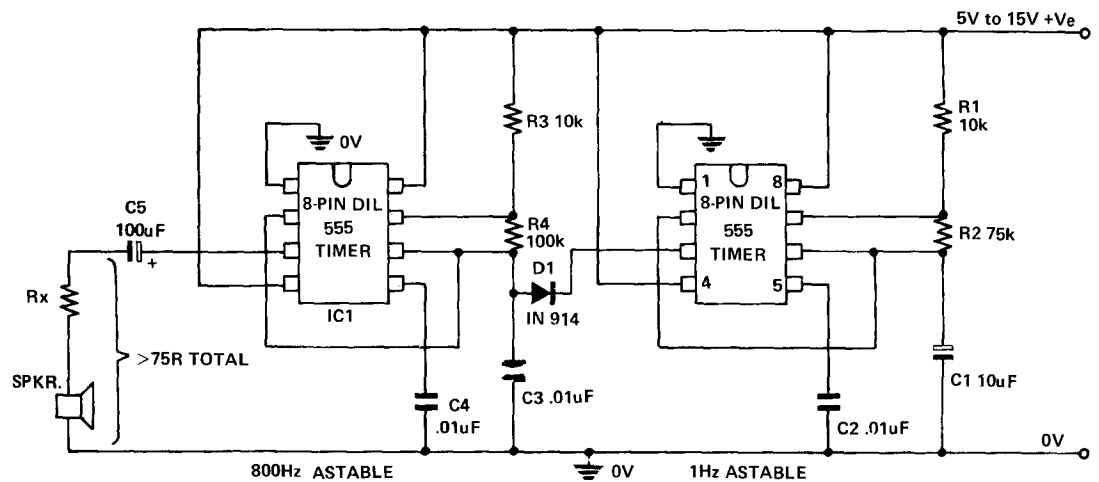temperature, and d) by over-temperature.

(a)

(b)

(c)

(d)

Fig.25. Pulsed-tone (800Hz)
alarm call generator.

give automatic dark-activated operation. Here, R4 and R5 are wired as a fixed potential divider that sets 1/2 Vcc on the emitter of Q1, LDR and R7 are wired as a light-sensitive potential divider that applies a variable voltage to the base of Q1, and the collector of Q1 is taken to RESET pin 4 of the IC, which is normally biased to ground via R6.

In use R7 is adjusted so that the voltage to the base of Q1 is greater than 1/2 Vcc under 'daylight' conditions, so Q1 is cut off, but under 'dark' conditions Q1 base is biased below 1/2 Vcc, so it is driven on, thus, under daylight conditions Q1 is cut off, so the 555 astable is disabled, with its output driven low, by 4.7 kΩ resistor R6 which is wired between pin 4 and ground. Under 'dark' conditions, on the other hand, Q1 is biased on, so pin 4 is positively biased, and the astable operates normally and activates the LED.

The LDR used in the above circuit can be any cadmium-sulphide photocell that presents a resistance in the approximate range 470 Ω to 10 kΩ under the minimum 'dark' turn-on condition.

The dotted section shows how the 555 astable circuit can be used as a 12 volt relay pulser, which turns the relay on and off at a rate of one cycle per second. The relay can be any type with a coil resistance greater than 60 Ω.

Alarm Generator

Figure 24 shows the connections for making an 800 Hz monotone alarm-call generator. The circuit can be used with any supply in the range 5 to 15 volts, and with any speaker impedance. Note, however, that Rx must be wired in series with speakers of less than 75Ω impedance, and must be chosen to give a total series impedance of at least 75 Ω, to keep the peak speaker currents within the 200 mA driving constraints of the 555. The available alarm output power of the circuit depends on the speaker impedance and supply voltage used, but may be as great as 750 mW when a 75 Ω speaker is used with a 15 volt supply.

The above circuit can be modified so that it is activated by darkness (a), by brightness (b), by an under-temperature (c), or by an over-temperature (d). Pin 4 is disconnected from the +Ve supply, and connected to the triggering circuit, which is designed around Q1. This works in the same way as already described for the automatic (dark-activated) LED flasher. The LDR used in the light-activated versions of this circuit can be any cadmium-sulphide photocells that present resistances in the approximate range 470 Ω to 10 kΩ at the desired turn-on levels. The thermistors used in the temperature-activated versions of the circuit can be any negative-temperature-coefficient types that present resistances in the same range at the required turn-on temperatures.

Alarms and Sirens

The next four diagrams show a variety of useful alarm-call generator circuits. The Fig 25 circuit generates an 800 Hz pulsed tone alarm call. Here, IC1 is wired as an 800 Hz alarm generator, and IC2 is wired as a 1 Hz astable which gates IC1 on and off via D1 once every second, thus causing a pulsed-tone output signal to be generated.

The Fig 26 circuit generates a warble-tone alarm signal that simulates the sound of a police siren. Here, IC1 is again wired as an alarm generator and IC2 is wired as 1 Hz astable multi-vibrator, but in this case the output of IC2 is used to frequency modulate IC1 via R5. The action is such that the output frequency of IC1 alternates symmetrically between 500 Hz and 440 Hz, taking one second to complete each alternating cycle.

The circuit of Fig 27 generates a 'wailing' alarm that simulates the sound of an American police siren. Here, IC2 is wired as a low frequency astable that has a cycling period of about six seconds. The slowly varying ramp waveform on C1 of this chip is fed to pnp emitter follower Q1, and is then used to frequency modulate alarm generator IC1 via R6. IC1 has a natural centre frequency of about 800 Hz. The circuit action is such that the alarm output signal starts at a low frequency, rises for 3 seconds to a high frequency, then falls during 3 seconds to the low frequency, and so on ad infinitum.

Finally, to complete this quartet of alarm generator circuits, the Fig 28 circuit generates a siren alarm signal that is a simulation of the 'Red Alert' alarm used in the STAR TREK T.V. programme: This signal starts at a low frequency, rises for about 1.15 seconds to a high frequency, ceases for about 0.35 seconds, then starts rising again from a low frequency, and so on add infinitum. The circuit action is as follows:

IC2 is wired as a non-symmetrical astable multivibrator, in which C1 alternately charges via R1 and D1, and discharges via R2, thus giving a rapidly rising and slowly falling 'sawtooth' waveform across C1. This waveform is fed to pnp emitter follower Q1, and is thence used to frequency modulate pin 5 of IC1 via R6. Now, the frequency modulation action of pin 5 of the IC1 astable circuit is such that a rising voltage on pin 5 causes the astable frequency to fall, and vice versa; consequently the sawtooth modulation signal on pin 5 causes the astable frequency to rise slowly during the falling part of the sawtooth and collapse rapidly during the rising part of the sawtooth. The rectangular pin 3 output of IC2 is used to gate IC1 off via npn common emitter amplifier Q2 during the collapsing part of the signal, so only the rising parts of the alarm signal are in fact heard, as in the case of the genuine STAR TREK 'Red Alert'.
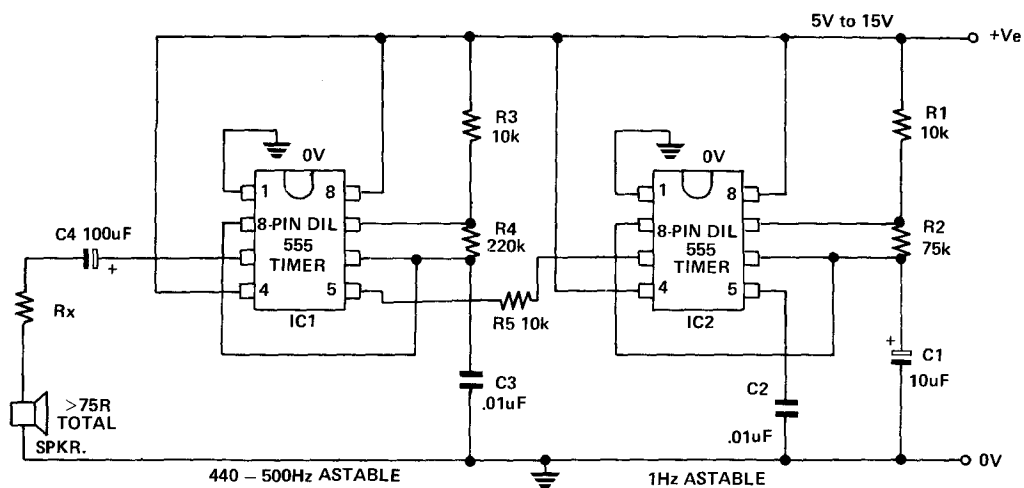
555 TIMER APPLICATIONS

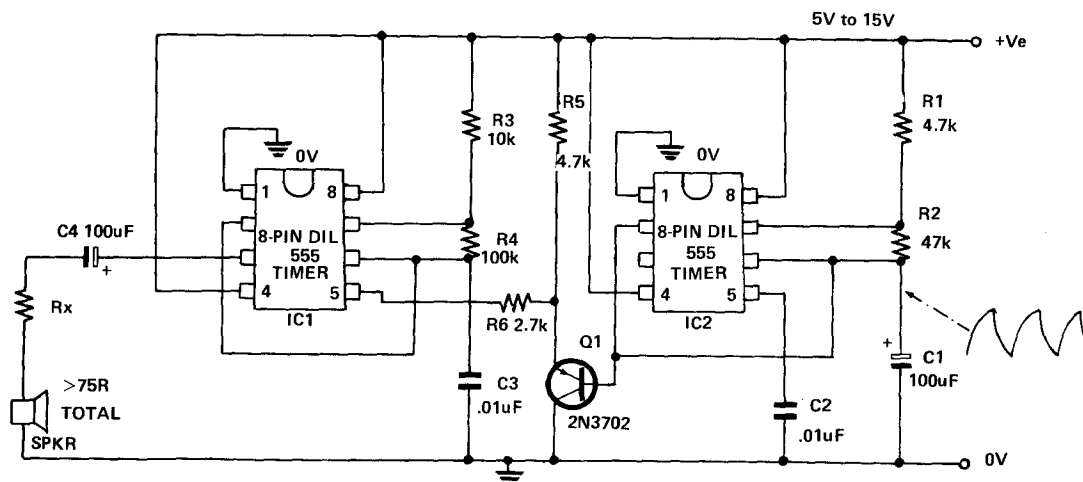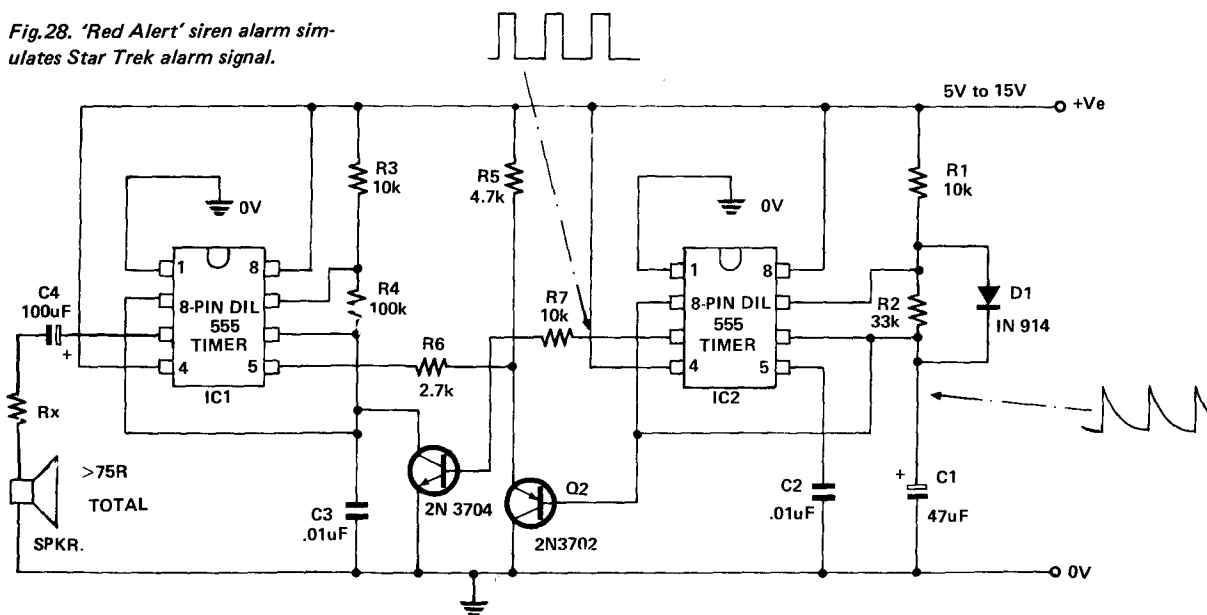Fig.26. Warble-tone alarm call generator simulates British police siren.

Fig.27. 'Wailing' alarm simulates American siren.

Fig.28. 'Red Alert' siren alarm simulates Star Trek alarm signal.

555 TIMER APPLICATIONS

DESCRIBED BY R. M. MARSTON

Miscellaneous Applications

To complete the story of the 555, this final section shows a miscellany of 555 applications, of varying degrees of usefulness. Figure 29 shows how a single 555 can be used as the basis of an event-failure alarm or a missing-pulse detector, which closes a relay or illuminates an LED if a normally recurrent event fails to take place.

The operating theory of the circuit is fairly simple. The 555 is wired as a normal monostable pulse generator, except that transistor Q1 is wired across timing capacitor C1 and has its base taken to TRIGGER pin 2 of the IC via R3: The TRIGGER pin is fed with a train of pulse or switch-derived clock input signals from the monitored event, and the values of R1 and C1 are selected so that the monostable period of the IC is slightly longer than the repetition period of the clock signal.

Thus, each time a clock pulse arrives, a monostable timing period is initiated via pin 2 of the IC, and C1 is discharged and the pin 3 output is driven high via transistor Q1. Before each monostable period can terminate, a new clock pulse arrives, and a new monostable period is initiated, so the pin 3 output terminal remains high so long as clock input pulses continue to arrive within the prescribed period limits. Should a clock pulse be missed, or the clock period exceed the pre-determined limits, however, the monostable period will be able to terminate normally, and pin 3 of the IC will go low and drive the relay or LED on. The circuit thus functions effectively as an event-failure alarm or missing-pulse detector. With the component values shown, the monostable has a natural period of about 30 seconds. This period can be varied via R1 and C4 to satisfy specific requirements.

Figure 30 shows how a couple of 555s can be used to make a pulse-width modulation (PWM) circuit. This circuit can be used for transmitting coded messages, or for applying variable power to a load at maximum efficiency.

Here, IC1 is wired as 1 kHz astable multivibrator, which is used to feed a continuous train of clock pulses to the pin 2 TRIGGER terminal of IC2, which is wired as a normal mono-
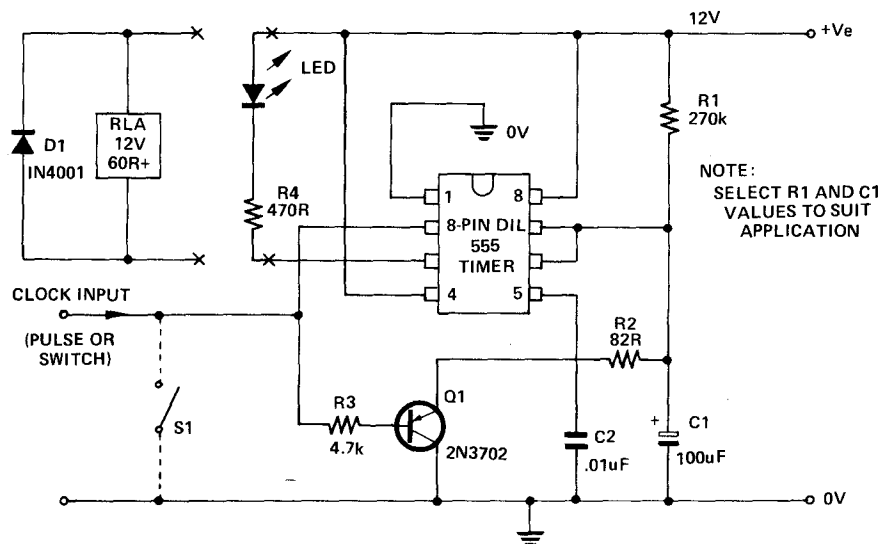

Fig. 29. Event-failure alarm or missing-pulse detector has relay or LED output.
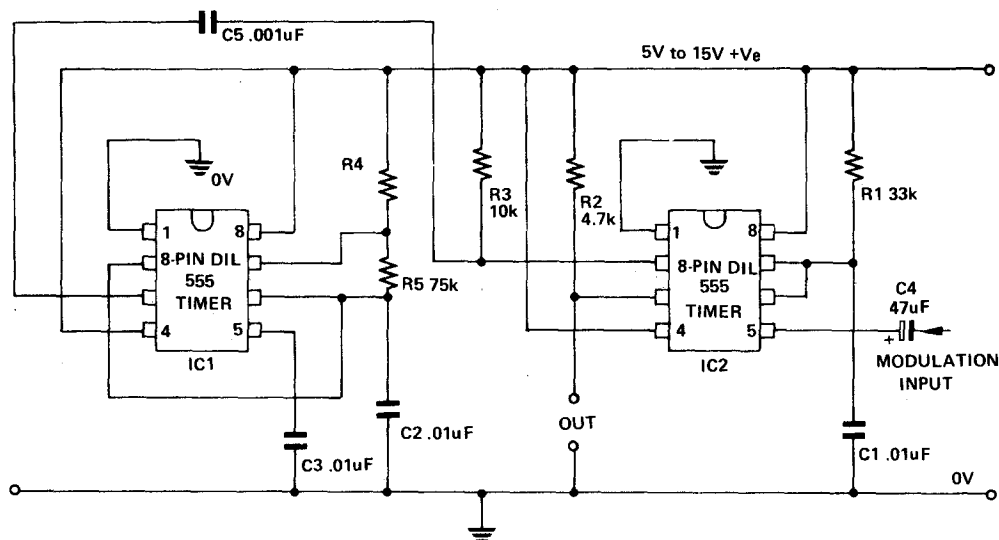
555 TIMER APPLICATIONS

Fig. 30. Pulse-width modulation
(PWM) circuit.

stable multivibrator or pulse generator and has a natural monostable period of approximately 0.36 mS. The external modulation signal is fed to the pin 5 CONTROL VOLTAGE terminal of the monostable via C4, and determines the instantaneous widths of the generated pulses. Thus, the circuit generates a train of pulse-width modulated (PWM) pulses at a fixed repitition frequency of 1 kHz.

Scope timebase

Figure 31 shows how a basic 555 monostable multivibrator can be modified so that it generates a linear ramp waveform of fixed duration each time it is triggered: The circuit can form the basis of an excellent oscilloscope time-base generator. The circuit works just like a normal monostable circuit, except that timing capacitor C1 is charged via constant-current generator Q1 during each timing cycle, thus causing a linear ramp voltage to be generated across C1.

When a capacitor is charged via a constant-current generator, the voltage across the capacitor rises linearly at a predictable rate that is determined by the magnitudes of the charging current and the capacitance. The relationship can be expressed as:

Volts-per-second = I/C, when I is expressed in amps and C is expressed in farads.

In this circuit the charging current can be varied over the approximate range 90 μA to 1 mA via R4, thus giving rates of rise on the .01 μF capacitor of 9 V-per-ms to 100 V-per-ms. Now, remembering that each monostable period of the 555 circuit terminates at the point when C1 voltage reaches 2/3 Vcc, and assuming that a 9 V supply is used (giving a 2/3 Vcc value of 6 V), it can be seen that the monostable cycles of the Fig 32 circuit have periods variable from 666 μS to 60 μS. Periods can be increased beyond these values by increasing the C1 value, or vice versa. Note when using this circuit that its supply rail must be stabilised if stable timing periods are to be obtained.

If the circuit of Fig 31 is to be used as the basis of an oscilloscope timebase, note that the input driving signal must first be converted to a good square wave, from which suitable trigger pulses can be derived via C3 and R5. The minimum useful ramp period that can be obtained from the circuit is about 5 μS, which, when expanded to give full deflection on a ten-division 'scope screen, gives a maximum timebase speed of 0.5 μS-per-division. Flyback beam-suppression signals can be
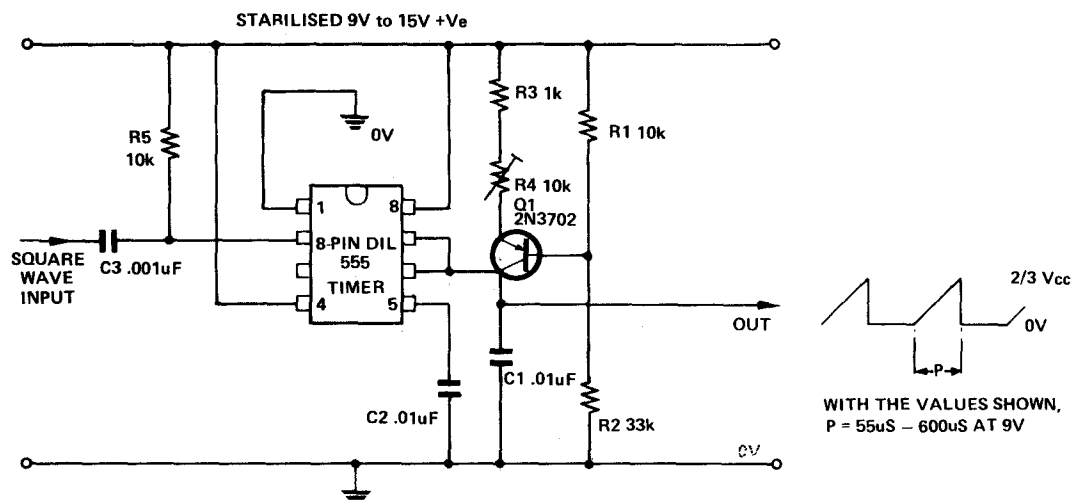

Fig. 31. Triggered linear-ramp
generator can be used as the
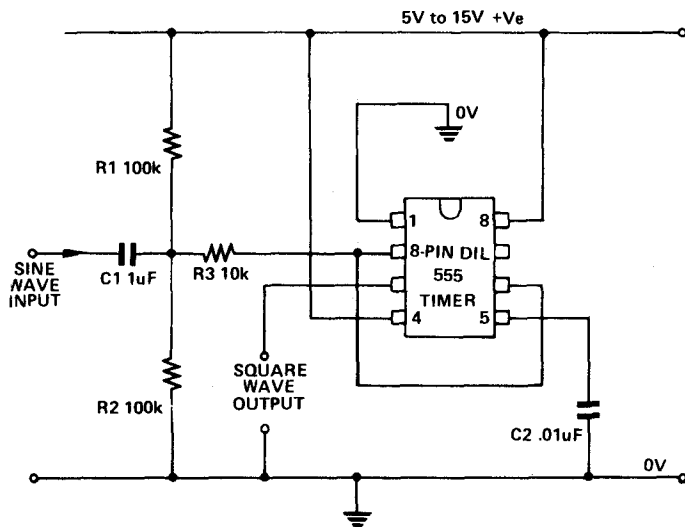basis of an oscilloscope time-
base.

Fig. 32. 555 Schmitt trigger circuit acts as excellent sine/square converter up to about 150 kHz.

Fig. 33. Dark-activated relay switch has built-in backlash.

derived from the pin 3 OUTPUT terminal of the IC.

The 'timebase' circuit gives superb signal synchronisation at trigger frequencies up to about 150 kHz. If the timebase is to be used with input signal frequencies greater than this, the input signals should be divided down via a single- or multi-decade digital divider. Using this technique, the timebase can be used to view input signals up to many MHz.

Figure 32 shows how a 555 can be connected for use as a simple but effective Schmitt trigger or Sine/Square converter. The circuit acts as a good converter at input frequencies up to 150 kHz or more. It works by changing its output state each time the pin 2 input signal swings from above the 2/3 Vcc level to below the 1/3 Vcc level, or vice versa. Resistor R3 is wired in series with pin 2 of the chip to ensure that the input signal is not adversely influenced by the transition action of the IC.

Figure 33 shows how the basic Schmitt circuit can be adapted to a dark-activated relay driving application by wiring light-dependent potential divider R1-LDR to the pin 2 input
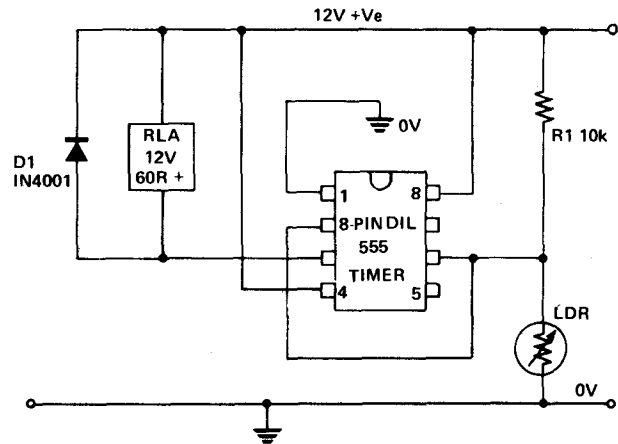
terminal of the IC. This circuit has an inherently high degree of input backlash, and is likely to be of value in only very specialised applications.

A far more useful relay-driving switching circuit is shown in Fig 34. This circuit has negligible input backlash, and can be used as either a light- or temperature-activated switch. In light-activated applications R1 is wired in series with a cadmium-sulphide photocell that presents a resistance in the approximate range 470 Ω to 10 kΩ at the required turn-on level. Dark-activated operation can be obtained by using the connections shown in Fig 34a or light-activated operation can be obtained by using the connections shown in Fig 34b.

For temperature-activated operation, R1 must be wired in series with a negative-temperature-coefficient thermistor, This thermistor must present a resistance in the range 470 Ω to 10 kΩ at the required turn-on level. Under-temperature operation can be obtained by using the connections shown in Fig 34c, or over-temperature operation can be obtained by using the connections shown in Fig 34d.

Fig. 34. Minimum-backlash relay switch can be activated by (a) dark, (b) light, (c) under-temperature, or (d) over-temperature.
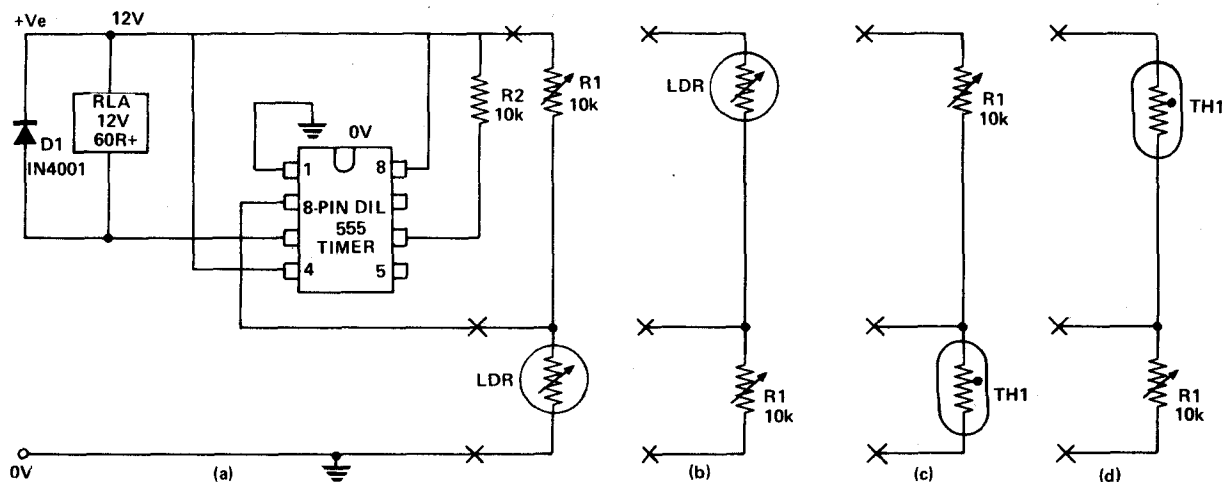
1 kHz Analogue Frequency Meter

This circuit needs a square-wave input driving signal with a peak-to-peak amplitude of two volts or greater. In this circuit the 555 is wired as a standard monostable multivibrator or pulse generator, and is powered from a regulated 6 V supply. Transistor Q1 is used to amplify the square wave input signals to a level suitable for triggering the monostable stage, and the output of the monostable is fed to 1 mA fsd meter M1 via multiplier resistor R5 and offset-cancelling diode D1. This meter gives a reading that is directly proportional to the frequency of the square wave input signals, and its operating theory is as follows:

Each time the monostable multivibrator is triggered it generates a pulse of fixed duration and fixed amplitude. If we assume that each generated pulse has a peak amplitude of 10 V and a period of 1 ms, and that the pulse generator is triggered at an input frequency of 500 Hz, it can be seen that the pulse is high (at 10 V) for 500 ms in each 1000 ms total period, and that the MEAN value of output voltage measured over this total period is 250 ms/1000 ms x 10 V = 5 V, or 50% of 10 V. Similarly, if the input frequency is 250 Hz the pulse is high for 250 ms in each 1000 ms total period, so the mean output voltage equals 250 ms/1000 ms x 10 V = 2.5 V, or 25% of 10 V. Thus, the mean value of output voltage of the pulse generator, measured over a reasonable total number of pulses, is directly proportional to the repetition frequency of the generator.

Normal moving coil meters are 'mean' reading instruments, and in the Fig 35 circuit a 1 mA fsd moving coil meter is wired in series with voltage multiplier resistor R5, which sets the meter sensitivity at about 3.4 V fsd, and is connected so that it reads the mean output voltage of the pulse generator. This meter thus gives a reading that is directly proportional to
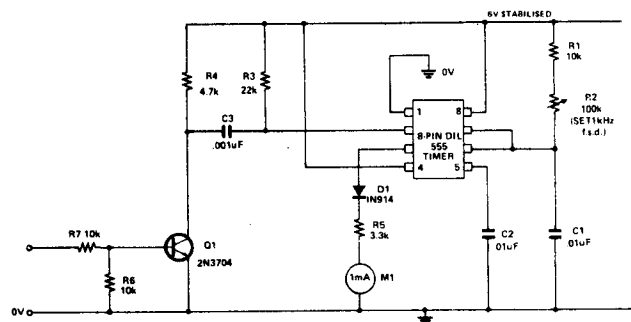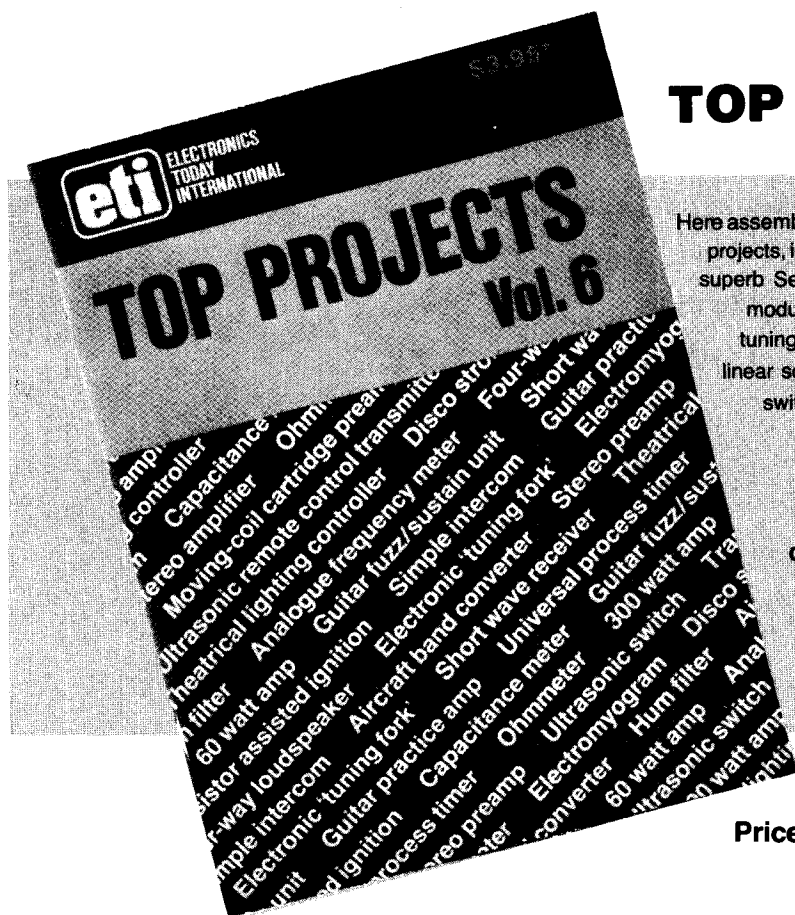

Fig. 35. Simple 1 kHz linear-scale analogue frequency meter.

frequency, and the circuit thus acts as a linear-scale analogue frequency meter. With the component values shown the circuit is intended to read fsd at 1 kHz. To set up the circuit initially, simply feed a 1 kHz square wave signal to its input, and then adjust R2 (which controls the pulse lengths) to give full-scale reading on the meter; all adjustments are then complete.

The full-scale frequency of the above circuit can be varied from about 100 Hz to about 100 kHz by suitable choice of C1 value. The circuit can be used to read frequencies up to tens of MHz by feeding the input signals to the monostable circuit via a single- or multi-decade digital divider, thereby reducing the input frequencies to values that can be read by the monostable circuit. The circuit can form the basis of an excellent and inexpensive multi-range linear-scale analogue frequency meter.
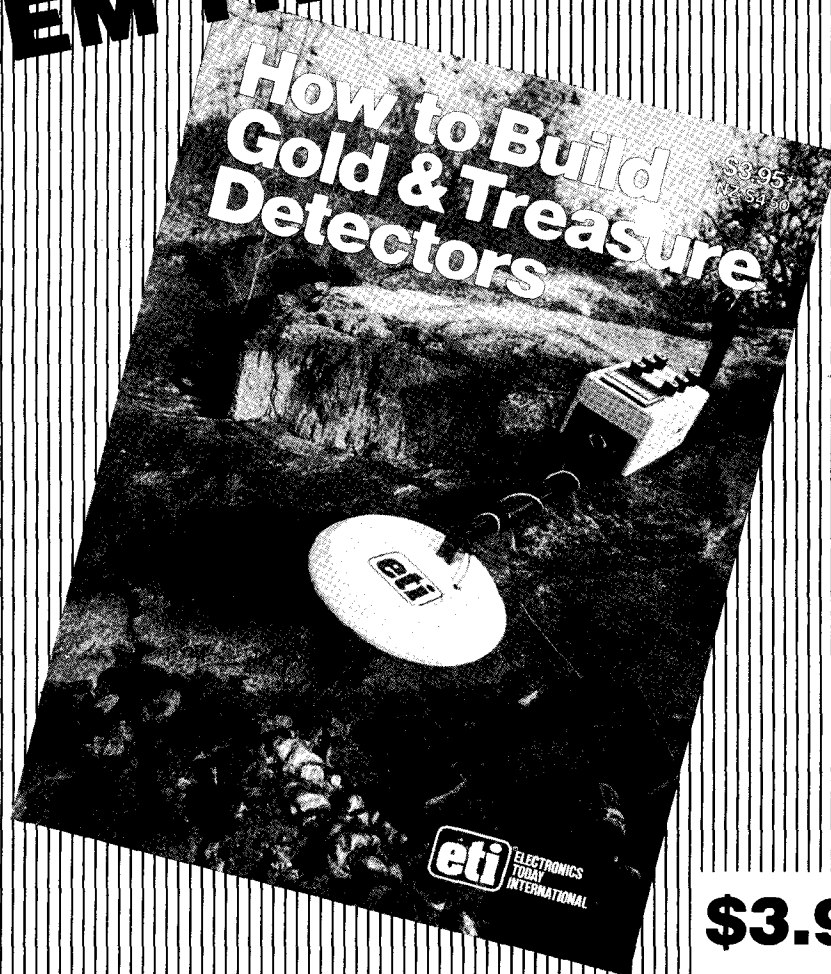
CMOS-a practical guide

Inherently rugged CMOS logic has many advantages over other logic families — high noise immunity and uncritical power requirements are but two. This new five-part series explains in practical down-to-earth terms just how CMOS should be used.

THE AVAILABILITY OF THE CD4000 series of chips brings CMOS to the forefront of logic technology to rival TTL in many applications. CMOS is far less critical as regards power supplies and possesses high noise immunity as well as capabilities which are not offered by other logic families.

In this series of articles we shall give various circuits which illustrate the use of CMOS. We suggest that these circuits can be 'breadboarded' on ETI Utilboard (Aug 1975 issue) using DIL sockets — there is nothing like trying it to see. Some of the circuitry, of course, is capable of realisation in a number of different logic families, so that, in these cases, we will merely be introducing equivalents of familiar devices. However, some of the applications we give show the major possibilities of CMOS.

HANDLING AND USE

Firstly we shall deal with the disadvantages of CMOS and get these behind us before we look more closely at some of the virtues. The first point is that these devices are very susceptible to

damage from surges of over-voltage from static electricity and unearthed test equipment. When you come to buy any CMOS ICs you should find them with their leads buried in foam. This foam is conductive and protects the device by putting all the pins together so do not remove it until the IC is to be put in circuit. If you run out of foam for storing devices then stick them into a piece of soft balsa wood. Whatever else you do, must not keep them in plastic containers or use ordinary plastic foam as this may develop a great deal of static. It is in fact a good rule to keep the devices away from all plastics as much as possible including any nylon clothing.

It is sensible to use IC sockets for the more expensive devices and also for any chip you may wish to re-use, but if you do solder them in solder the V_{DD} pin first, then V_{SS} and then all the others. The reason for this is that the common ranges of CMOS have internal protection devices which operate fully only when the supply lines are connected. While we are on the subject of soldering, check that your iron and any other instruments you may use

(meters, oscilloscopes, etc) are all properly earthed.

The only other real disadvantages of CMOS compared with TTL are that it is slower (typical gate rise time 25 nS) and that a few operating precautions are necessary.

Firstly, all unused inputs must go somewhere. The alternatives are tying unused inputs to used inputs, either to the supply line as appropriate, or to a supply line via a resistor (220 kΩ is usually about right). The latter solution is particularly helpful for inputs to which off-board connections are to be made. This avoids leaving the input "floating" until it is wired in. The other point is to ensure that the chips do not have signals at their inputs when the power supply is not on.

Now we shall consider a few of the advantages of CMOS. Most of these will come out more clearly later and so we shall just mention them briefly here.

The principal virtue is the ease of choice of power supply. This may be anywhere between three and fifteen volts at low current. The actual power required depends on operating frequency (see Fig 1) being comparable with TTL at ten MHz but it is in the region of a few microwatts at sub-kilohertz speeds. Voltage regulation is not requied but operating speed and current consumption rise with increasing supply voltage. For most practical purposes CMOS will run off a nine volt battery or the simplest of mains power supplies. Other advantages include high noise immunity and analogue possibilities. Before we proceed with some circuitry the table of operating conditions (Table 1) should be studied. The limits shown in the table should be adhered to rigidly.

```
TABLE 1 CMOS OPERATING LIMITS          CD400A SERIES
STORAGE TEMPERATURE ......................... −65 to +150°C
OPERATING TEMPERATURE ....................... −40 to +85°C
SUPPLY VOLTAGE LIMITS (V_DD−V_SS) ............... −0.5 to +15 V
PACKAGE DISSIPATION ............................ 200 mW max.
INPUT VOLTAGE ............................... V_SS ⩽ Vin ⩽ V_DD
RECOMMENDED SUPPLY VOLTAGE (V_DD−V_SS) ....... +3 V to +15 V

Unused inputs should be tied to a supply line. No input should be present
when the supply lines are off.
```
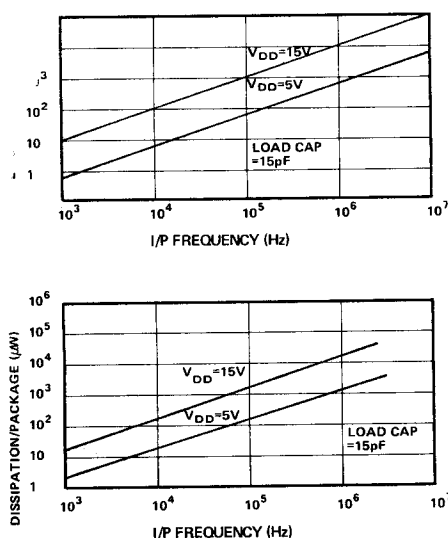
Fig. 1. Power dissipation in CMOS as a function of frequency for a) a simple gate and b) an MSJ package.

SIMPLE GATES

It is an unpleasant fact that it seems one must always start considering any subject at its least interesting parts and it is hardly surprising that the least interesting logic ICs are the simple gates.

We shall assume that the reader is familiar with truth tables and terminology and logic consequently our discussion will mainly be on the subject of monostable and astable multivibrators. For ease of future reference a list of basic CMOS gates and their pin-outs is given in Fig. 2. It is worth remembering that inverters may be realised by tying together all the inputs of a NAND or NOR gate, thus allowing a circuit requiring two NOR gates and two inverters to be constructed for a single type 4001A package.

ASTABLE MULTIVIBRATORS

The basic CMOS astable is shown in Fig. 3. This could of course be built using any of the packages in Fig 2 with the exceptions of the 4030A and 4050A, indeed, the 4049A could produce three of these circuits simultaneously. The period is approximately 1.4RC (R in ohms, C in farads), and the waveform may have a non-unity markspace ratio due to the voltage at which the inverters switch (called the transfer voltage — V_{tr}) not being exactly half way between V_{DD} and V_{SS}. The frequency is also dependent on the supply voltage. In

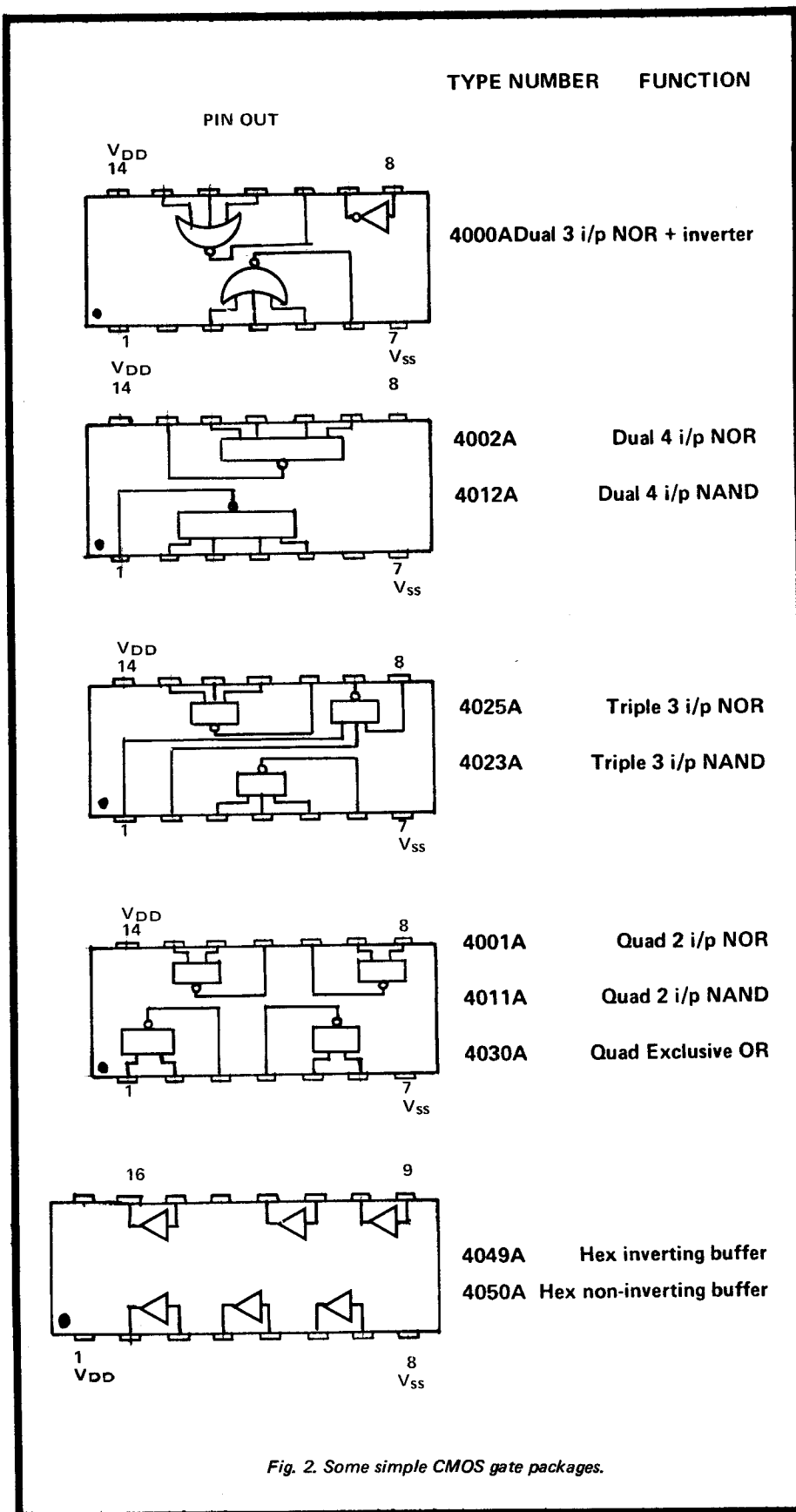| TYPE NUMBER	FUNCTION
4000A	Dual 3 i/p NOR + inverter
4002A	Dual 4 i/p NOR
4012A	Dual 4 i/p NAND
4025A	Triple 3 i/p NOR
4023A	Triple 3 i/p NAND
4001A	Quad 2 i/p NOR
4011A	Quad 2 i/p NAND
4030A	Quad Exclusive OR
4049A	Hex inverting buffer
4050A	Hex non-inverting buffer

Fig. 2. Some simple CMOS gate packages.

CMOS - a practical guide
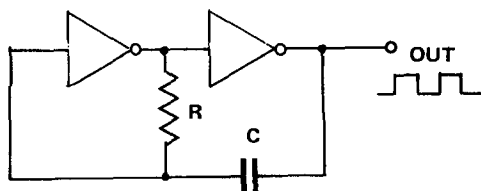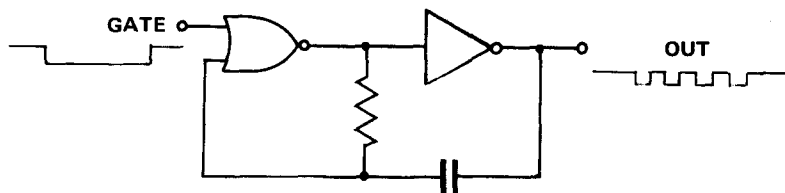

Fig. 3. Basic CMOS as table.

Fig. 4. Improved astable multivibrator.

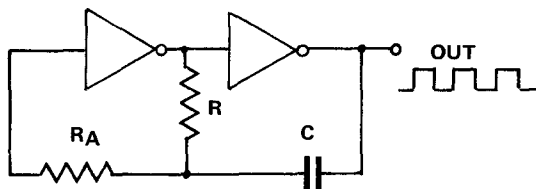Fig. 5. Duty cycle adjustment.

Fig. 6. A gated astable multivibrator.
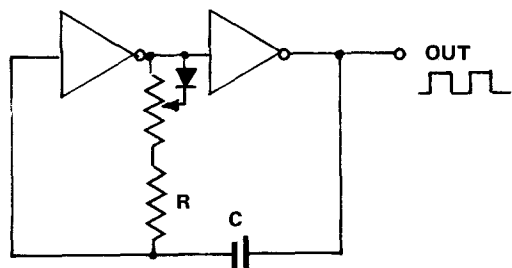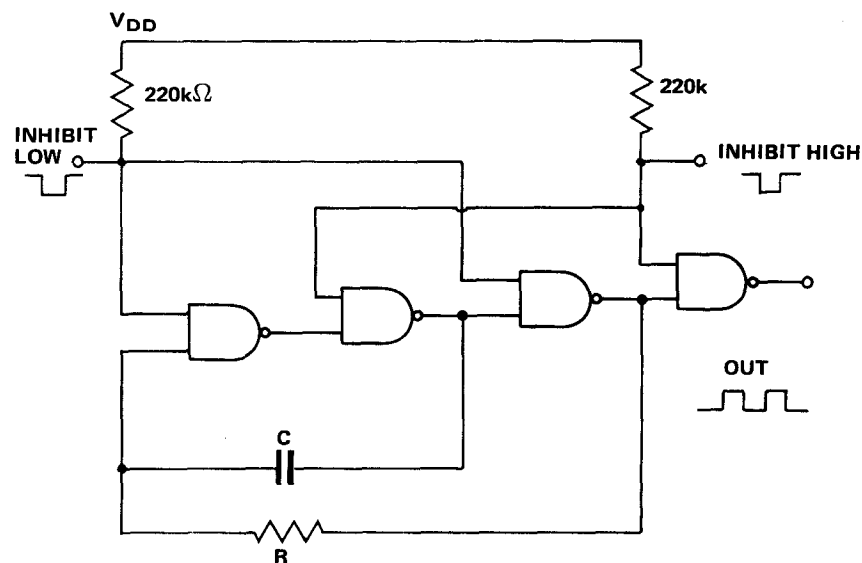

Fig. 7. This jitter-free multivibrator (with inhibits) may be made from a 4011A.
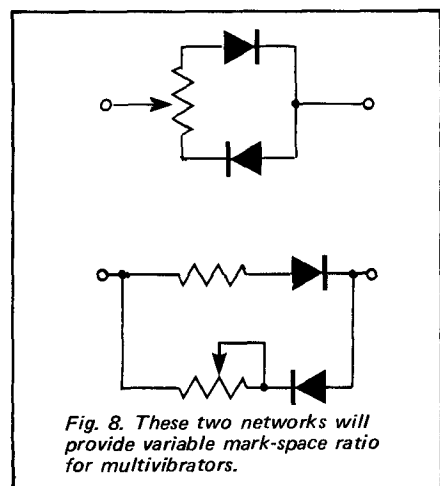

Fig. 8. These two networks will provide variable mark-space ratio for multivibrators.

keeping with normal practice, connections of the device to the supply voltage have not been shown.

The next few circuits will rectify some of the aberrations of the simple version. The addition in Fig. 4 of R_A, which should be at least twice as large as R, makes the frequency almost independent of the supply voltage over a wide range. The frequency of any of the circuits may be made variable by making R a variable resistor.

Duty cycle adjustment may be achieved using the circuit in Fig. 5 Altering the duty cycle will affect the frequency and the diode may have to be reversed to achieve the desired result.

GATING

A gated multivibrator is shown in Fig. 6 where the oscillator only runs when the gate input is low, thus producing "bursts" of output in synchronism with the control signal. Using a NAND gate instead of the NOR would cause the circuit to run when the gate was high instead of low.

One of the huge advantages of CMOS is the very high input impedance. As a consequence timing resistors can be very large — values in the hundred megohm region with capacitors of several microfarads can be a practical proposition.

Before we leave the astable multivibrator for a time we shall give one more circuit which corrects a tendency of all the preceding ones to "jitter" near the switching point. This requires an extra inverter and a fourth

has been added as an output buffer. There are also two inhibit inputs which stop the circuit with the output high or low, depending on which is used. The theoretical diagram is shown in Fig 7. Another feature of this circuit, and indeed virtually all the others, is that the timing resistor may be substituted by one of the networks in Fig 8 to give a variable mark-space ratio. They work because the diodes effectively change the value of the timing resistor depending on whether the capacitor is charging or discharging, ratios as large as 5000:1 maybe used.

MONOSTABLE MULTIVIBRATORS

The basic CMOS monostable is shown in Fig. 9. It is triggered by the input pulse's leading edge and produces a positive going output pulse. The period may vary by more than ±50% with different devices due to the dependence of the circuit on the transfer voltage of the inverter.

The circuit in Fig. 10 operates in an interesting way. The quiescent state is with the first and second inverter outputs at "0" and "1" respectively.

falling edge of the triggering pulse
es the first inverter go high, C2
ges through the diode up to VDD
the second inverter goes low thus
ating the output pulse. C1 recharges
ough R1 and crosses the transfer
oltage of the first inverter which
onsequently goes low and is isolated
rom C2 by the now reverse-biased
diode. Capacitor C2 then discharges
through R2 and causes the second
inverter to revert to its initial state, thus
completing the output pulse. The
advantage of all this is that inverters
fabricated on the same chip have similar
transfer voltages and if the two time
constants (R1C1 and R2C2) are made
identical, errors cancel out and the
period becomes well defined. It is in
fact approximately equal to 1.4R1C1 (=
1.4R2C2) and this circuit is capable of
being retriggered during the output
pulse.

Our last two monostables (Figs.11/12)
are non-retriggerable and the two
time constants should be made the same
as in the previous circuit. Figure 12 is
particularly interesting because the
circuit isolates the trigger input during
the output pulse as the charge on C2
holds one input to the NOR gate high,
thus keeping the output low
independently of the state of the trigger
input.

FREQUENCY DOUBLER

The frequency doubler shown in Fig.
13 works by differentiating the leading
and trailing edges of the waveform and
applying the resulting pulses to the two
inputs of a NAND gate. This produces a
complete output pulse at both the rise
and fall of the input signal. The values
of the discrete components will depend
on the desired frequency of operation.

THE 4007

The next device we are going to
consider has no equivalent in other logic
systems. It is described as a "dual
complementary pair plus inverter" and
its type number is 4007. It can perform
several different functions and while we
are discussing it we shall present a
number of useful circuits and have the
added advantage of learning a little
about the internal operation of CMOS.

In CMOS there are two different types
of field-effect transistors, namely
n-channel and p-channel enhancement
mode devices (see Fig. 14). What all this
means is that when based in the
conventional manner (drain positive in
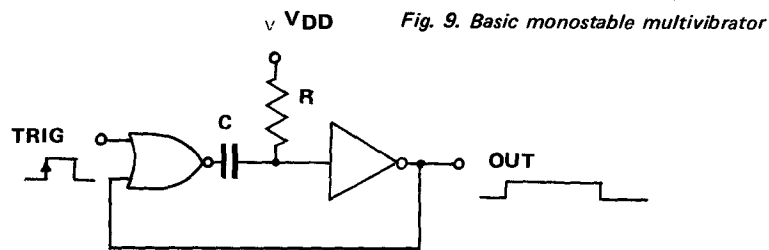n-channel devices but negative in
p-channel devices), the n-type turns on
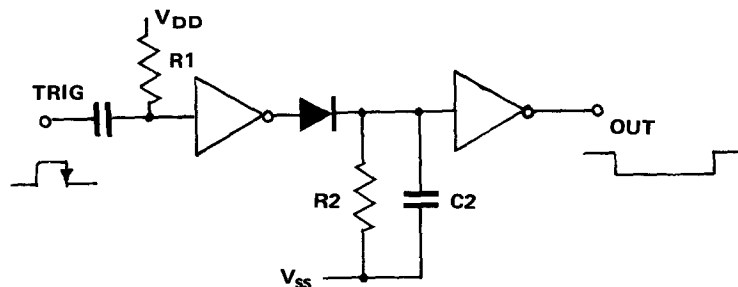

Fig. 9. Basic monostable multivibrator

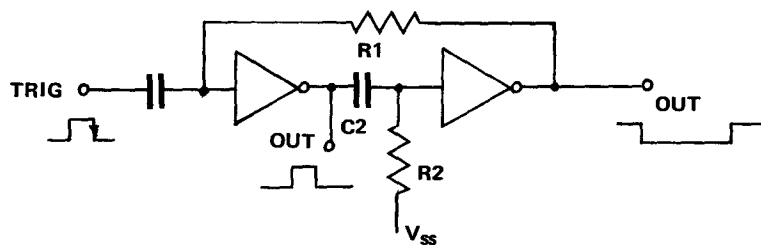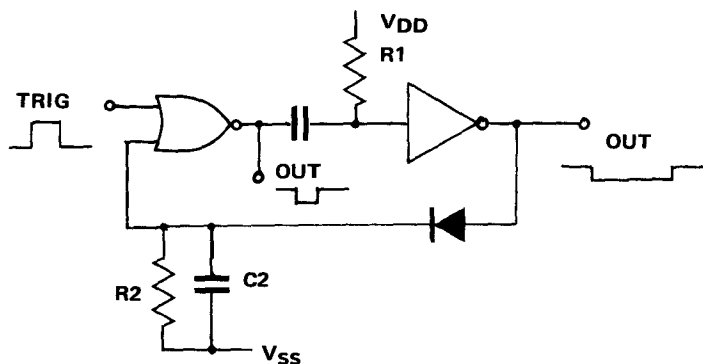Fig. 10. Improved monostable, R1C1-R1C1.

Fig. 11. A negative edge triggered monostable.

Fig. 12. Positive edge triggered monostable.

CMOS - a practical guide
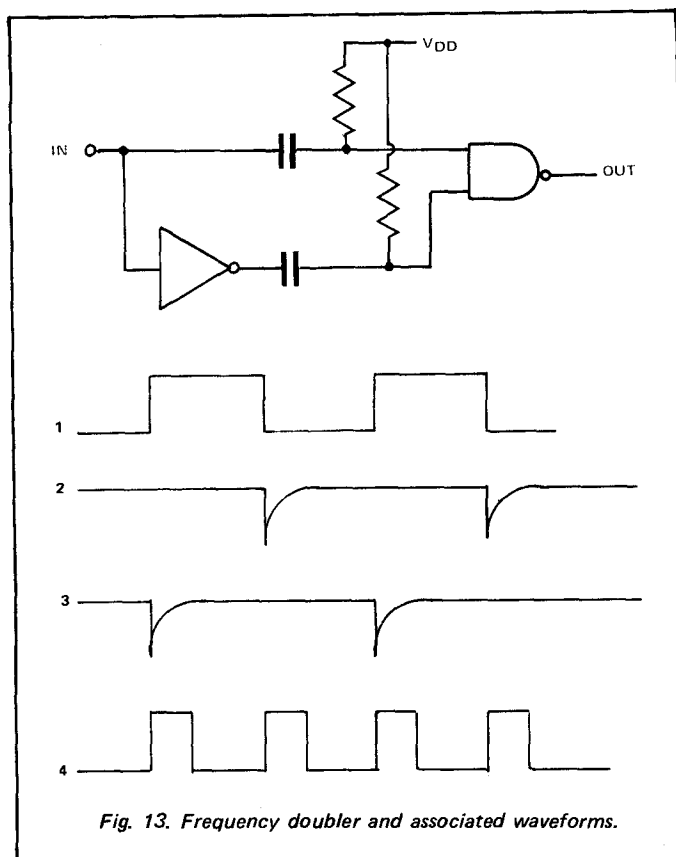

Fig. 13. Frequency doubler and associated waveforms.

Fig. 14. The 2 different MOSFETS used in CMOS, (a) n-channel, and (b) p-channel.

Fig. 15. CMOS inverter, $V_{SS} = 0V$. $V_{DD} = +3-15$ V.

when the gate becomes sufficiently positive with respect to the source and the p-type when it is sufficiently negative. A "turned on" device may be considered to have a resistance of the order of 500 ohms — 1k between source and drain whereas the equivalent resistance when "off" is about $10^9 \Omega$ The resistance at the gate is always very high ($> 10^{12}\Omega$) regardless of the state of the device.

The working of the CMOS inverter (Fig. 15) should now be fairly clear. When the input is "high" the bottom FET is turned on and the top one off. Thus the output voltage is held very low. When the input is low the FETs reverse roles and the output is high. Now look at Fig. 16 which shows the internal circuitry of the 4007. You should be able to see how joining a few pins together will allow three separate inverters to be produced. Reference to Fig. 17 will reveal how several other gates may be produced and their mode of operation should be relatively easy to discern.

TRANSMISSION GATES

There is another way of connecting two FETS which produces a result unique to CMOS. This is the transmission gate (Fig. 18). Here, due to the inverter, both FETS are either on or off simultaneously. When they are on, the path between input and output (they are interchangeable) may be regarded as a resistor of about 500 ohms — 1kΩ whereas when they are off the equivalent value is about 1000 MΩ.

Thus the device behaves as a switch capable of passing analogue signals with very little distortion provided that the load resistance is fairly high (≈ 100 kΩ). We shall have more to say about these "bilateral switches" later but while we are dealing with the 4007 Fig. 19 shows how to connect one as a single pole-double throw switch which will pass analogue signals in both directions.

Any of the three or less inverter circuits we have mentioned to date may be realised with a 4007, as may several more interesting designs. Figure 22 shows a linear frequency-to-voltage converter which works by charging a capacitor once for every input cycle, the
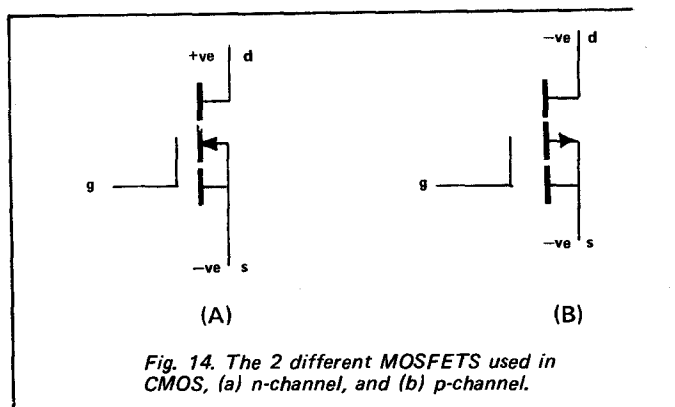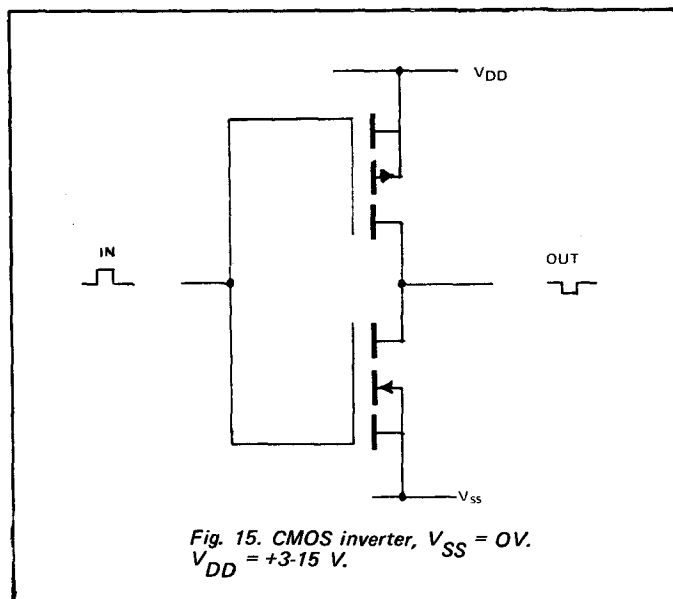
charge to do so being passed by a MOSFET into a summing amplifier. The component values given are based on an approximate five volt output for the given frequency. The resistor R1 should be made a 100 kΩ preset if it is required to set a range exactly. The capacitor C2 "smooths" the output and need not be changed from 10μF if fast response on the upper ranges is not needed. The linearity achieved on the top range will depend on the particular "741" used and if reliable operation is required higher speed op-amp should be used.

Figure 21 shows an alternative monostable multivibrator. We have already given a number of multivibrator circuits and so we shall say nothing more about this one except that it has an extremely small power consumption. This is due to the feedback connection (pins 12-6) which turns off the n-channel MOSFET during the discharge of the time constant. This circuit is also
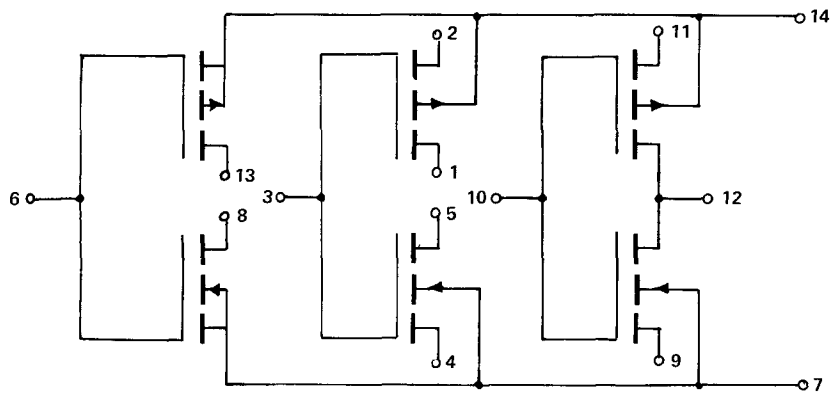
Fig. 16. The internal circuitry of the 4007 dual complementary pair plus inverter.

CONTROL IN

(1,5,12), (2,9); (11,4);
(8,13,10); (6,3)
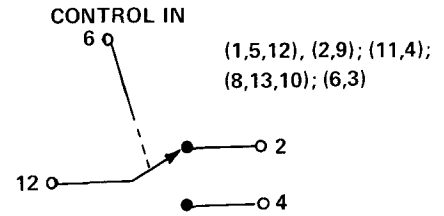
Fig. 18. The transmission gate.

TRIPLE INVERTERS
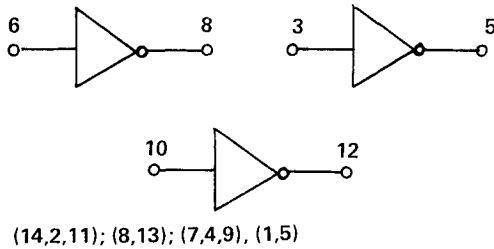
(14,2,11); (8,13); (7,4,9), (1,5)

THREE I/P NAND
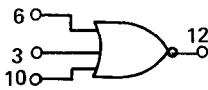
(1,12,13); (4,8); (2,14,11); (5,9)

CURRENT DRIVER

(13,8,1,5,12); (6,3,10); (14,2,11); (7,4,9)

Fig. 17. Some uses of the 4007;
pins bracketed should be connected together.
V_{DD} (pin 14) V_{SS} (pin 7)

THREE I/P NOR

(13,2); (12,5,8); (1,11); (7,4,9)
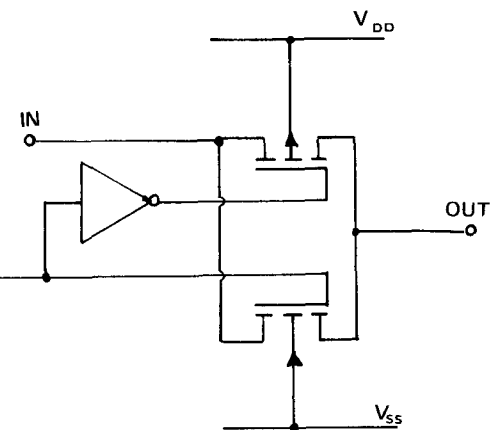
Fig. 19. An SPDT Electronic switch
which could be used for multiplexing.
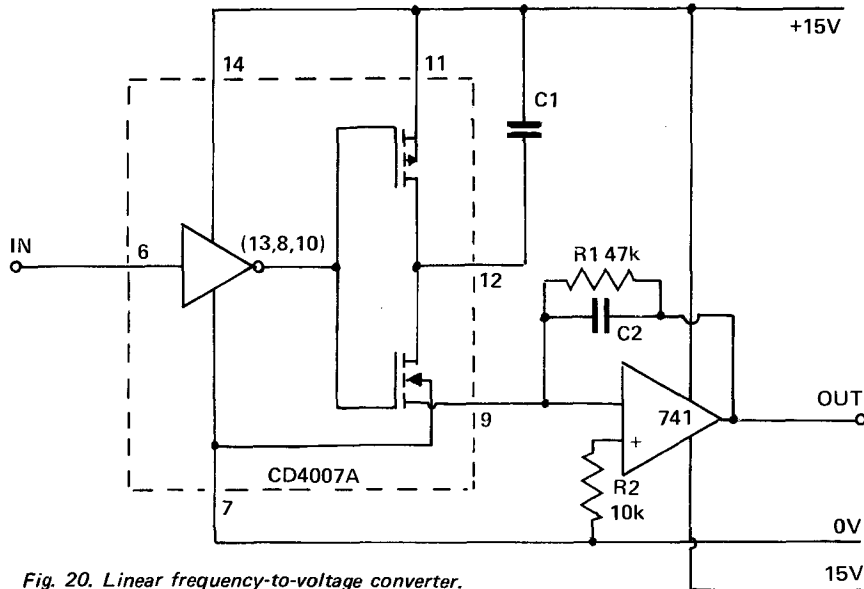
Fig. 20. Linear frequency-to-voltage converter.

FSD FREQUENCY	C1	C2
10Hz	1μF	10μF
100 Hz	100nF	1μF
1kHz	10nF	100nF
10kHz	1nF	10nF
100kHz	100pF	1nF

CMOS - a practical guide
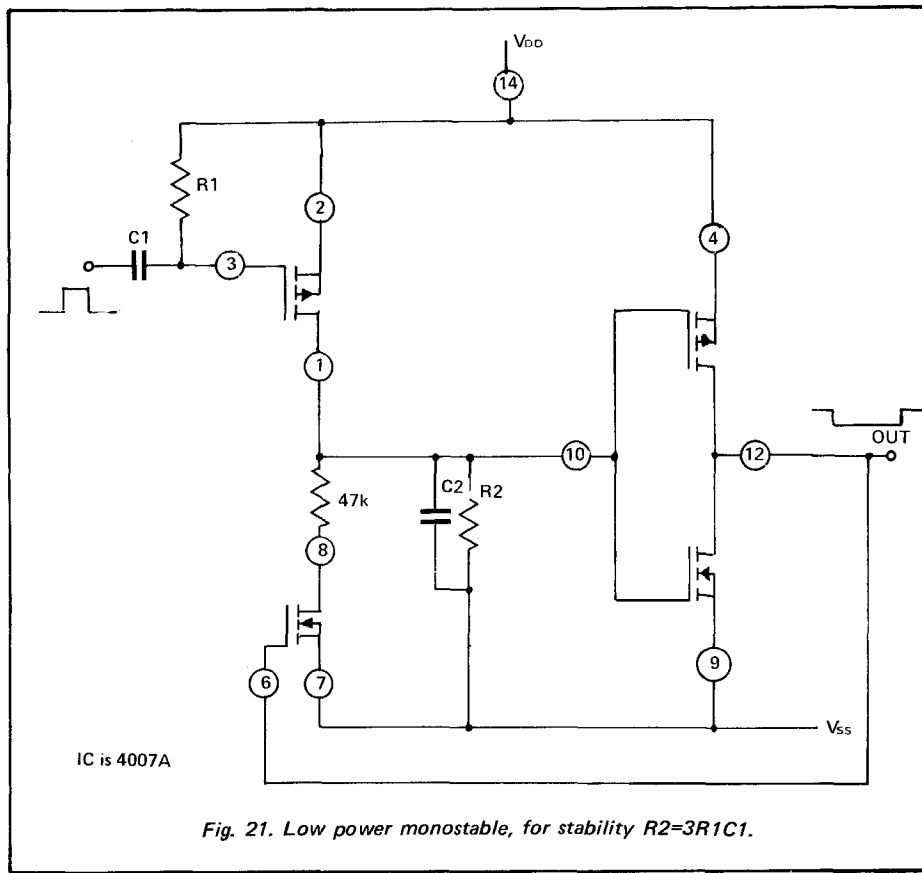
Fig. 21. Low power monostable, for stability R2=3R1C1.

Fig. 24. (a) Adjustable M−S ratio narrow range controller, (b) Exclusive−OR phase comparator.

an interesting demonstration of the use of components in the 4007 as discrete transistors.

A WIDE RANGE VCO

The voltage controlled oscillator shown in Fig. 24 uses two inverters as well as a separate transistor as a voltage controlled resistor. The inverters function as an astable multivibrator in the manner of Fig. 4 but the timing resistance is the parallel combination of R_T and the FET. As V_c varies between V_{dd} and V_{ss} so the resistance of the

FET varies between about 1 k and 1000 MΩ. If the upper value is limited to 10 M Ω by making R_t that value, then the circuit will sweep over a 10000:1 range in frequency. There

Fig. 22. A voltage controlled oscillator using the 4007 1C.

Fig. 23. Outline of PLL receiver.

Fig. 25. Exclusive—OR gate used as a conditional inverter together with the device's truth table.

IN-1	IN-2	OUT
0	0	0
0	1	1
1	0	1
1	1	0
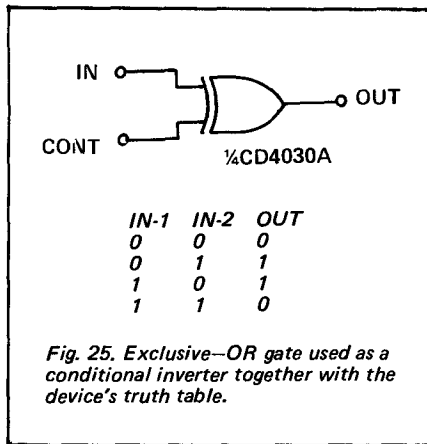
would seem to be scope here for experimenting with a pulse frequency modulation communications system. One might produce an analogue system although distortion would probably be high due to mismatching. The transmitter could be the circuit in Fig. 22 and the receiver a phase locked loop along similar lines (Fig. 23) using some sort of phase comparator and a low-pass R-C filter.

Fig. 26. Exclusive—OR gates driving liquid crystals segments, represented by resistors.

EXCLUSIVE-OR GATES

Exclusive-or gates, for example the 4030 (see Fig. 2), will function as phase comparators but they require a unity mark-space ratio to be effective. Perhaps a voltage controlled oscillator might be designed with a narrower range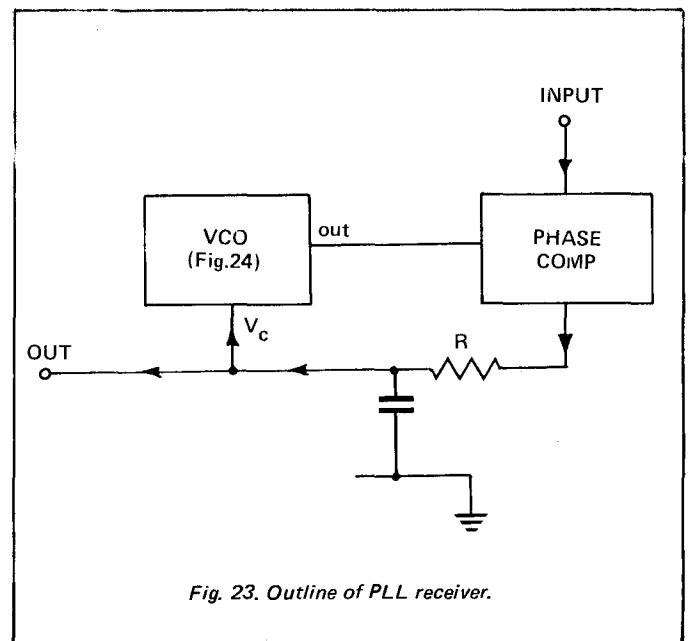 along the lines of Fig. 24 for both transmitter and receiver, together with a phase comparator and low pass filter as shown in Fig. 24. While we are on the subject of the exclusive-or function we shall consider two more uses of these devices. Figure 25 shows the exclusive-or truth table and its use as a conditional

inverter. This configuration causes the input signal to be inverted when the control input is high but not when it is low.

Liquid crystal displays are undoubtedly the readout devices of the future but they last longer in general if ac drive is used. If then a square wave is applied to one end of a liquid crystal segment and also to the other connection via a conditional inverter (see Fig. 26) then the control input will decide whether or not there is a net voltage across the segment.

CMOS and liquid crystal make an ideal combination for ultra-low power logic and display systems and so manufacturers have produced BCD-to-seven-segment decoders and drivers specifically for this application.

Their type numbers are 4054/5/6, the variations being due to the addition of latches and other refinements. These devices have too limited an appeal to justify a full description here.

COMMON SEMICONDUCTOR PIN-OUTS
(Cut this out, attach to strong card and stick it to your workshop wall)
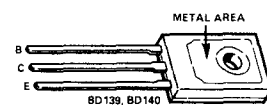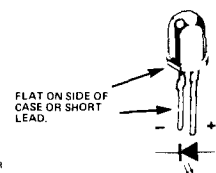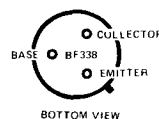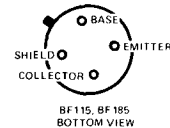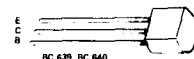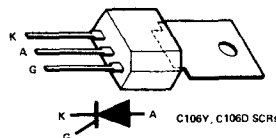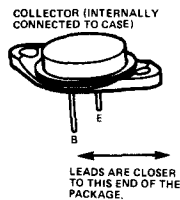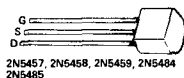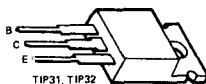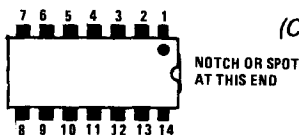
CMOS·a practical guide

Inherently rugged, CMOS logic has many advantages over other logic families — high noise immunity and uncritical power requirements are but two. This, the second article in this four part series deals primarily with quad switches.

Fig. 1 (a) Pin diagram of 4016A.

Fig. 2. DPDT electronic switch realised using a single 4016 and an inverter.
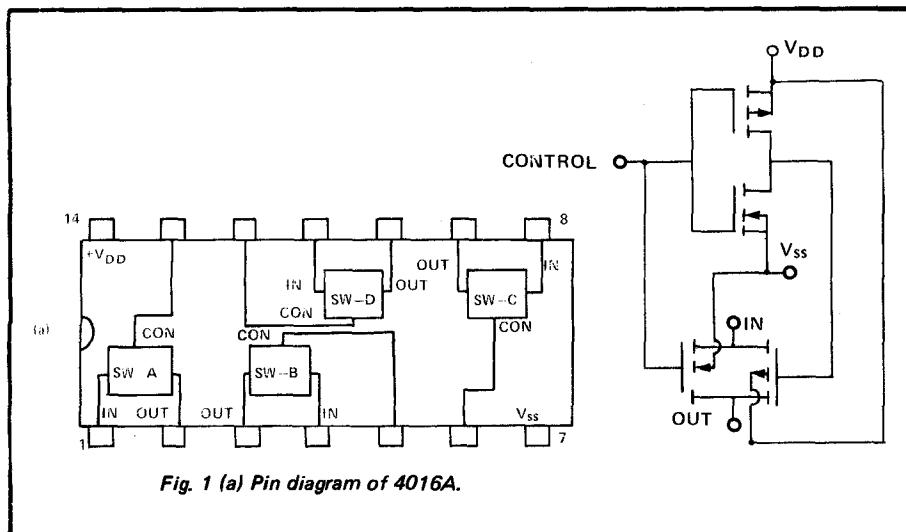
THE 4016A QUAD BILATERAL switch consists of four transmission gates of the type discussed in the previous article, each with its own control input. Each switch also has a signal input and output (although these are interchangeable). When the control input is held high the input to output path behaves like a pure resistance of about 300Ω but when it is low the equivalent value is of the order of $10^9\Omega$ at low frequency, even with fairly low supply voltages. It is impossible to give all the data which might be necessary for diverse applications — data sheets from a manufacturers may be required by the more adventurous experimenters. In any case the pin diagram in Fig. 1 should now be self explanatory.

It should be appreciated that the output impedance of the switch is fairly high and so for low signal distortion, a load greater than 10 kΩ is necessary. Using a high supply voltage (10-15 V) also helps to achieve this end. The gates will pass signals above the 10 MHz mark but as the frequency becomes higher, crosstalk between the switches (and distortion) will inevitably increase. It should be fairly clear how complicated switching systems may be realised but Fig. 2 has been included to guide constructors along the right lines.

ANALOGUE APPLICATIONS

Many uses of this device in audio equipment have already appeared in constructional articles in this magazine and so it is to two slightly less obvious applications that we shall turn now. Figure 3 shows a sample and hold unit. When the control input is high the output tracks the input, but when it goes low the output remains frozen at the value it was at the instant of transition. The operation of the circuit

Fig. 3. A simple sample and hold circuit.

Fig. 4. An analogue delay line (a) basic circuit and (b) composition of the voltage followers used.

Fig. 5. Digital component control using the 4016A (a) digital variable resistor (b) digital variable capacitor.

n = DIGITAL CONTROL INPUTS

CMOS·a practical guide

DIGITAL INPUTS
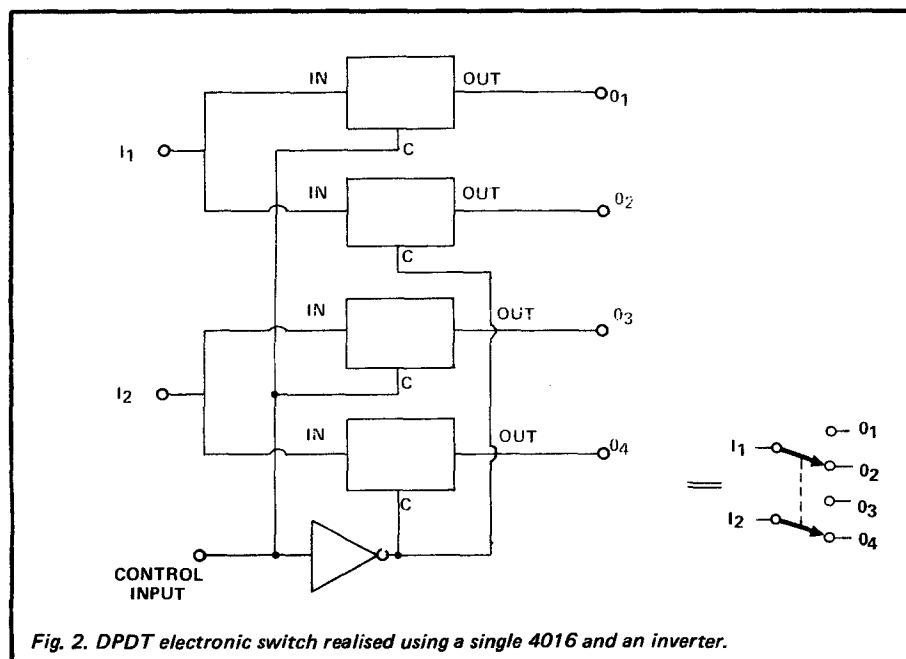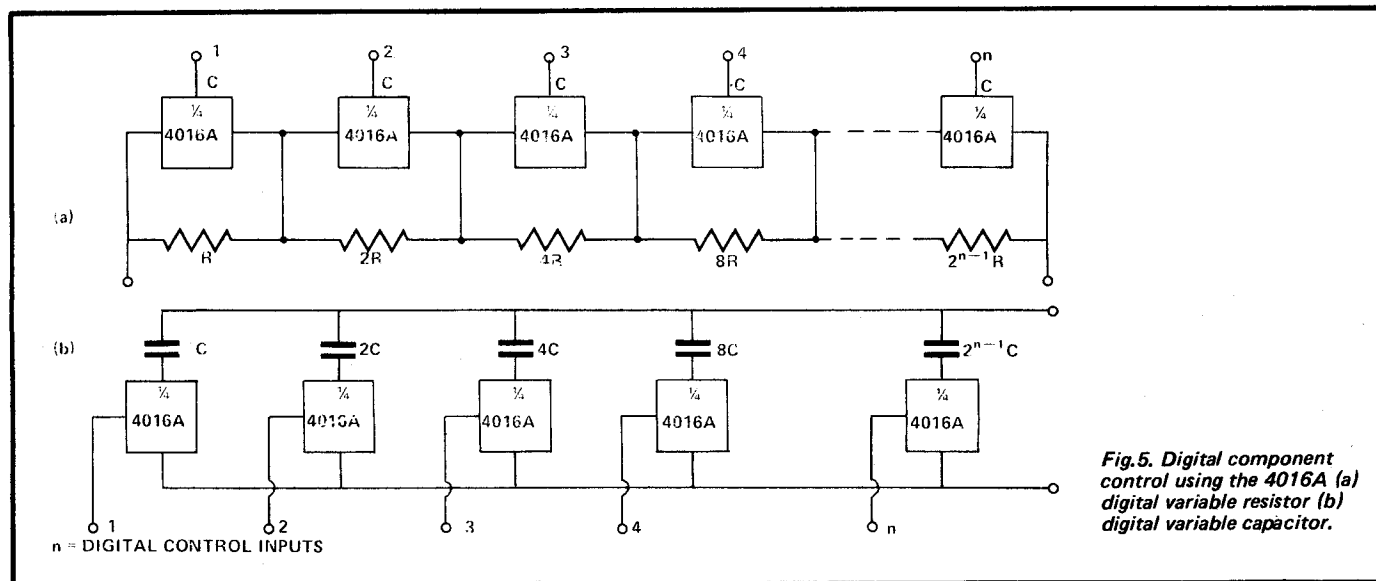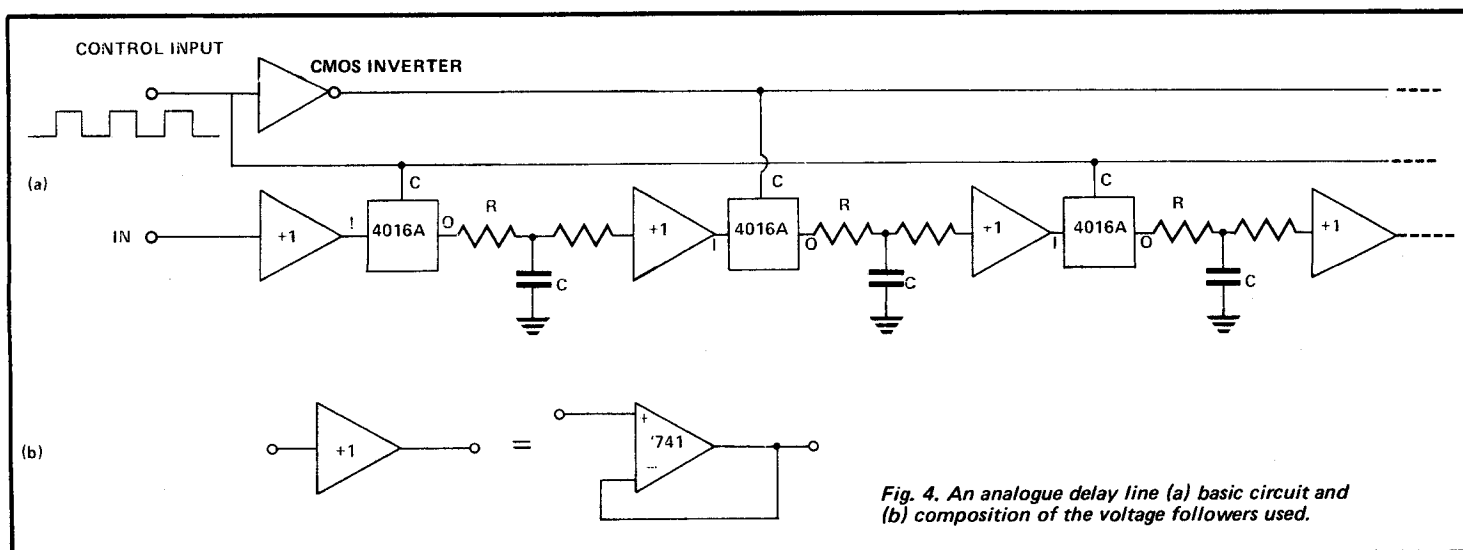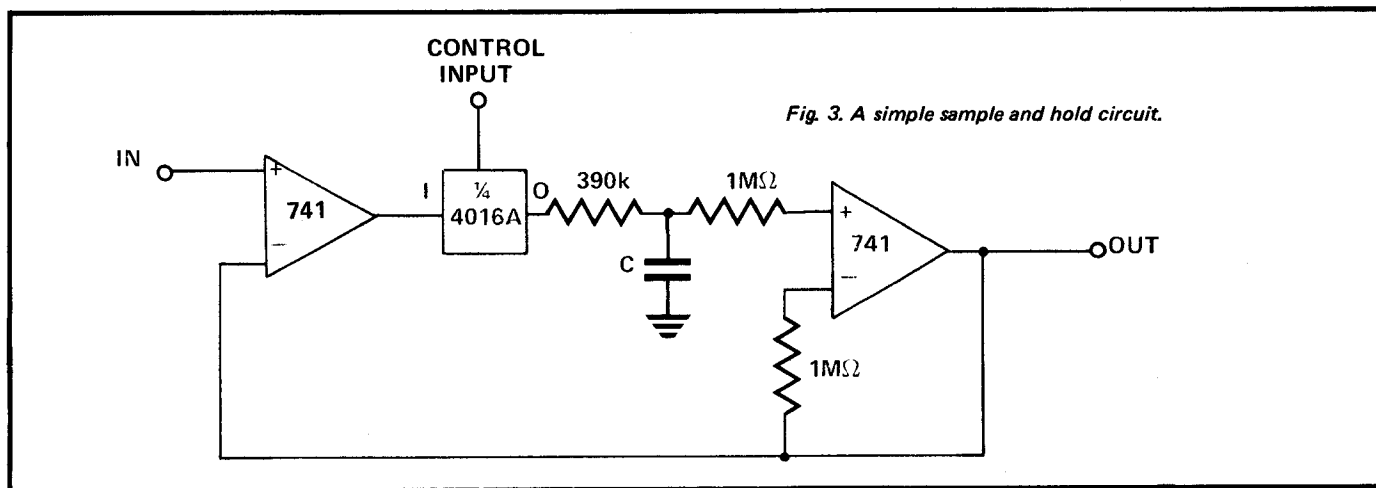
Fig. 6. Basic voltage fed digital to analogue converter.

DIGITAL INPUTS

Fig. 7. The R-2R ladder network D-A converter.

is generally self evident and it may be regarded as two voltage followers, one consisting of two op-amps with the *output following the input*, the other is just the second op-amp which follows the voltage stored on the capacitor. It is advisable to take care with the layout, as with all op-amp circuits, due to the huge open loop gain of these devices. The value chosen for C is a compromise between "slewing rate," that is the rate at which the circuit tracks a sudden change of input, and "holding ability" which is the length of time the circuit will hold a signal without unreasonable decay. To give some sort of guide, for a 10 kHz square wave to the control input, a 0.01µF capacitor seems to

optimise the performance. The value of the resistors is also worth experimenting with.

An extension of the sample and hold concept is the analogue delay line wich is shown in its basic form in Fig. 4. The sequence of amplifiers and gates can be extended to any desired length to achieve a longer delay, the only limitation being that in extreme cases the control lines may need to be buffered. It should be observed that alternate stages of the circuit are driven by an identical clock waveform and so the circuit works by shifting the voltage on alternate capacitors during alternate clock phases. The value of the passive components and the clock frequency

will have to be optimised for specific applications, low frequencies give long delays but high distortion.

DIGITAL COMPONENT SELECTION

There are a few fairly straight-forward uses of the 4016A in digital component selection which we will mention here because, in certain fields, they are very useful. Figure 6 shows how to produce digitally controlled resistance and capacitance networks wh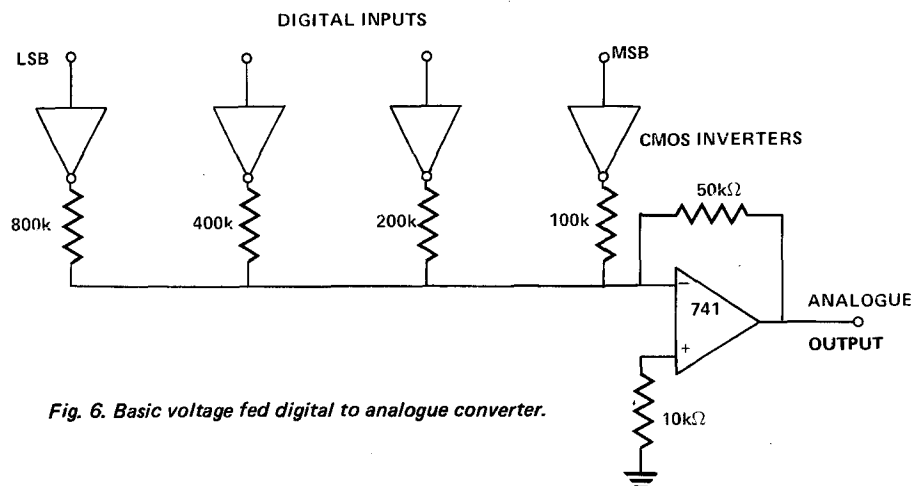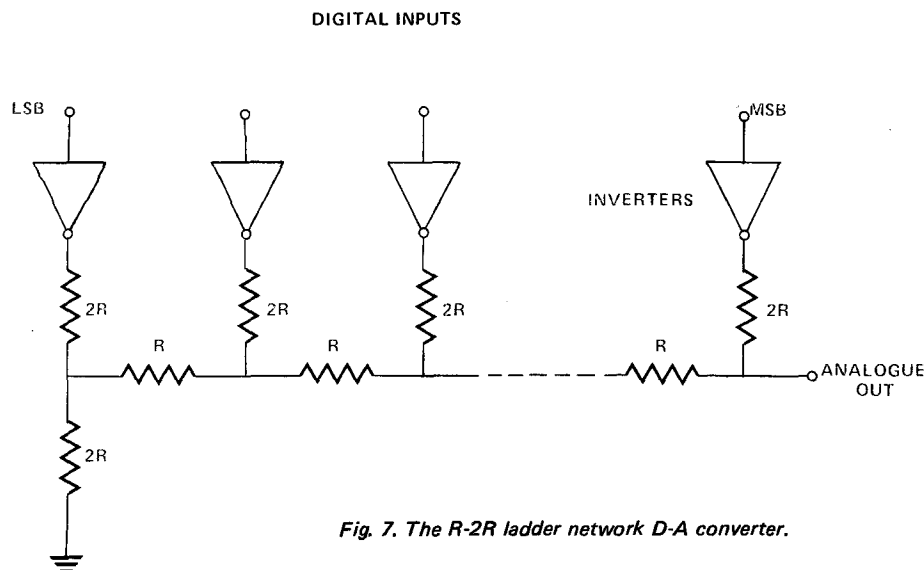ich will vary the magnitude of the quantity in question from its basic value up to $2^n - 1$ times that amount, where "n" is the number of gates and binary control bits. The resistor network can be used to produce a digitally gain controlled amplifier by placing it in the feedback loop of an op-amp and this can be used as a staircase generator — as well as to produce more interesting waveforms. One application of the digital capacitor is to produce a digitally controlled sweep generator by using it as the capacitor in one of the multivibrators we discussed last month.

Any type of component may be switched in and out of circuit by the 4016A. One possibility that is useful in some circumstances is to use the information on filter design in "Electronics — it's easy" to produce digital filters of different descriptions. The main thing to remember when using all these ideas is that the impedance of the component that is being switched must, at the desired frequency of operation, be large compared to the 300Ω of the 4016A gates.

D-A AND A-D TECHNIQUES

The next use of the quad bilateral switch that we are going to consider is digital to analogue conversion, but first we are going to look at the subject of conversion on its own. Figure 6 shows a D-A converter of the type known as a weighted resistor network. The working of this circuit is easy to see and the reason it is practical in so simple a form in CMOS is that the high and low output levels from a simple gate are within a few millivolts of the supply line thus providing accurate voltage levels to the summing resistors. This simple form has disadvantages, one of which will become immediately apparent if you consider the diversity of resistor values required for an eight or nine bit version. It is for this reason that the R-2R network shown in Fig. 7 is more popular for most applications. It is difficult to explain how this configuration works without becoming
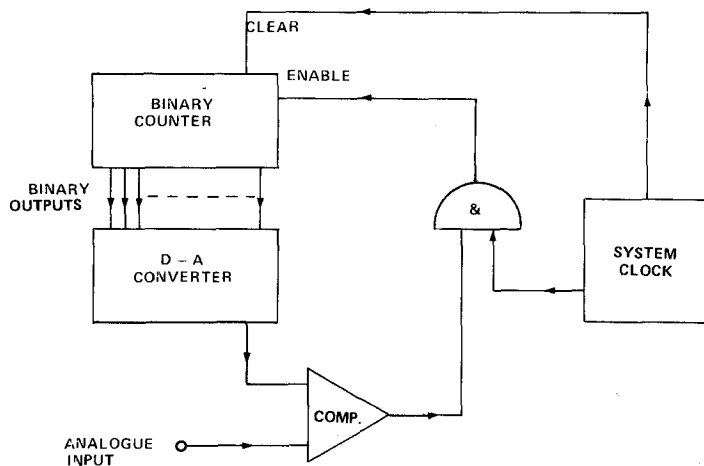
Fig. 8. Outline of A-D converter using 'Binary ramp' method. The binary counter counts up to the required value so that the output is correct during only part of the cycle.
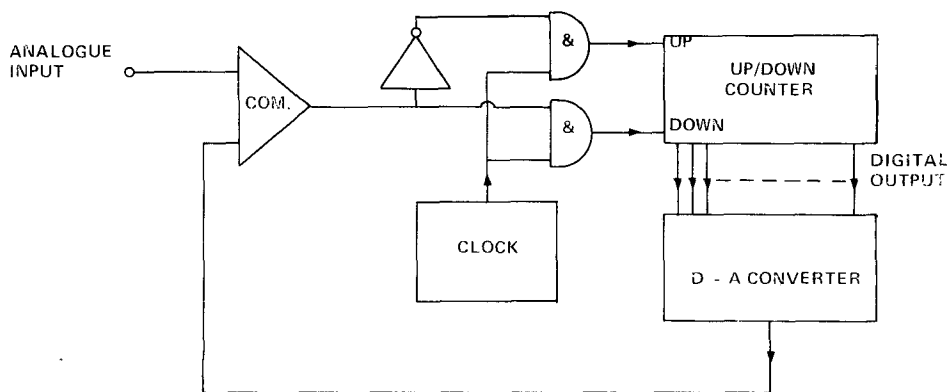
Fig. 9. A continuous counter type A-D converter. The clock just prevents the system oscillating at high frequency.

involved in mathematics but it becomes fairly clear if a three bit version is written down with the voltages in and analysed.

The basic idea for one sort of A-D converter or encoder as they are often called is shown in the block diagram of Fig. 8. This variety is called a binary ramp encoder and is similar to the ramp integration method often encountered in modified form in digital measuring equipment. The principle of operation of our binary version is that a binary counter counts up from zero until it reaches the equivalent of the analogue input. Contrast this approach with the continuous counter method outlined in Fig. 9. This provides a continuous approximation to the input and is thus generally more useful for continuously monitoring a single channel of information. It would be an advantage in many cases to have the counters working in BCD for ease of readout but this leads to complications. Finally on this subject it should be pointed out that these are circuits for experimentation and are unlikely to be directly applicable to any specific situation. They have been included because of the ease with which they may be realised in CMOS compared, say, to TTL.

A-D MULTIPLIER

As far as digital to analogue conversion is concerned (using the 4016A) we can take the idea a little further. What in fact we do (Fig. 10) is to use an arbitrary analogue voltage to feed the
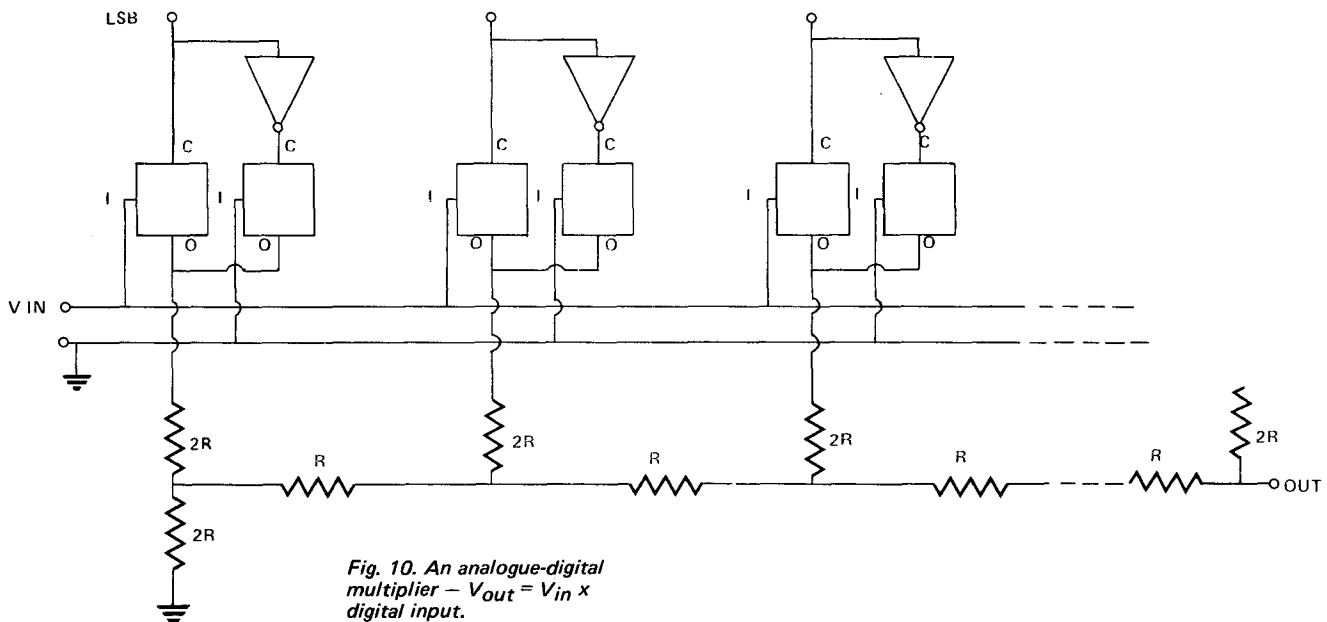
Fig. 10. An analogue-digital multiplier — $V_{out} = V_{in} \times$ digital input.

CMOS - a practical guide

resistor ladder and so we multiply this input by the digital input and produce an analogue result. This "hybrid multiplier" is an interesting circuit, particularly because the analogue input voltage may be ac thus producing several interesting waveforms and, on a more serious note, it may find application in hybrid computing experiments. We shall now leave the 4016A having, it is hoped, suggested some of the slightly less obvious uses of this versatile IC.

Our next subject is flip-flops — and we shall assume that the reader is familiar with the working of these devices, and so the discussion will begin with the pinout diagrams in Fig. 11. The first two are standard, dual edge — triggered devices with "D" and J-K" type data inputs respectively. No doubt it is known that the "D" variety will divide the input frequency by two if "Q" is connected to "D" whereas the "J-K" type toggle, as this behaviour is called, when both "J" and "K" are held high. The set and reset inputs operate asynchronously (ie. independently of the clock) forcing the device into the "Q" = 1 and "Q" = 0 states respectivley. These inputs operate when taken high in contrast to most TTL because TTL inputs rest high when disconnected whereas CMOS inputs must never be allowed to "float" anyway. Both the 4016A and the 4027A will operate up to about 8 MHz.

The last device in Fig 11 (the 4042A) is a quad data latch of the sort often used for temporary storage of BCD digits in applications like frequency meter displays. If the polarity input is held low then the "Q" output follows the "D" input in each latch when the clock is also low. But on the rising edge of the clock pulse the outputs are isolated and retain the data present at that moment. When the polarity input is high all this works the other way round. The clock inputs to all these devices should have rise times of 5μs or less (at V_{dd} = 10 V).

Flip-flops on their own have uses in control circuitry and counters. If you wish to produce a counter to count through an odd sequence (a Gray code for example) it is advisable to find out about Karnaugh maps and associated techniques which aid the design process considerably. The standard form for such counters is a sequence of flip-flops whose inputs are derived from the outputs of the others by a few simple logic gates. As far as simple binary is concerned the required set-up is shown in Fig. 12 but we shall have a lot more to say on the subject of counters in general later.

The other main application of flip-flop is in shift registers. A shift register is a sequence of flip-flops so interconnected (see Fig 13) that on a clock pulse the content of each device is transferred to the next one down the line. The register so formed is referred to as a static device because, unlike some MOS devices available, data is not lost if it is not shifted for some length of time. One modification to the basic device is to provide inputs and outputs to individual flip-flops in the chain and in this form shift registers have many applications in serial-to-parallel and parallel-to-serial data conversion. This though is another subject which must wait until a little later in our discussions. ●

4013A DUAL D FLIP FLOP — VDD=14, Vss=7

4027A DUAL J–K FLIP FLOP — VDD=16, Vss=8

4042A QUAD D–LATCH — CLOCK = 5, POLARITY = 6, VDD =16, Vss = 8

BEFORE CLOCK PULSE		AFTER	
CLK	"D"	Q	Q̄
↑	0	0	1
↑	1	1	0

BEFORE CLOCK PULSE			AFTER	
CLK	J	K	Q	Q̄
↑	0	0	*	*
↑	0	1	0	1
↑	1	0	1	0
↑	1	1	†	†

* NO CHANGE
† CHANGES STATE

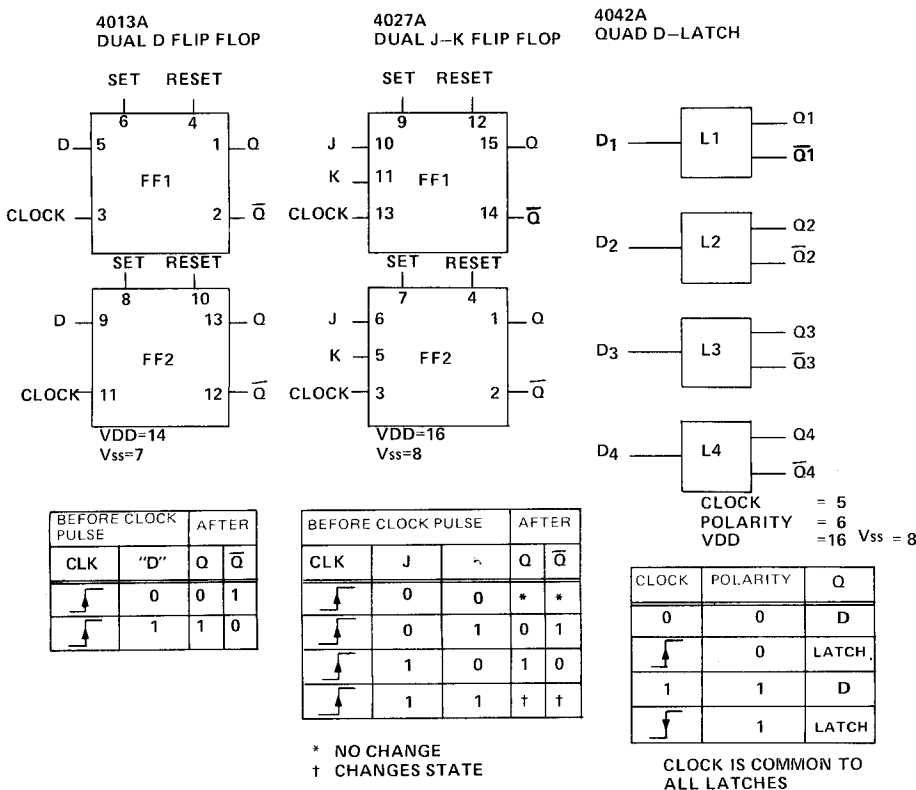| CLOCK	POLARITY	Q
0	0	D
↑	0	LATCH
1	1	D
↑	1	LATCH

CLOCK IS COMMON TO ALL LATCHES
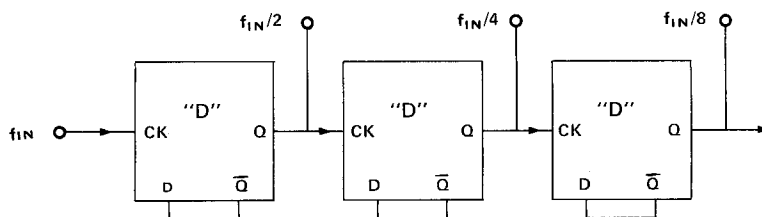
Fig.11. Three common flip-flop devices.

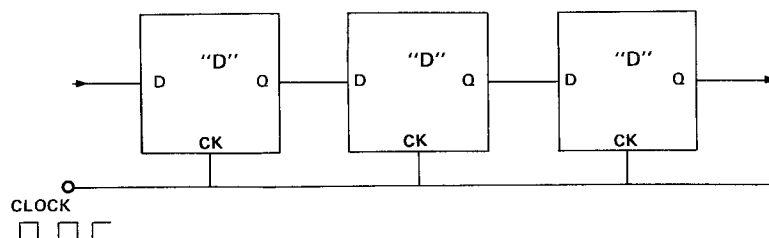Fig.12. Basic binary counting chain used for frequency division.

Fig. 13. Principle of shift register. The content of each "cell" is shifted one place to the right on the rising edge of the clock pulse.

CMOS-a practical guide

Inherently rugged, CMOS logic has many advantages over other logic families — high noise immunity and uncritical power requirements are but two. This, the third article in this series deals with counters.

OUR MAIN SUBJECT in this article is counters. It might well be true to say that the range available (compared to TTL) reflects the advances which have been made in other branches of electronics, particularly display technology. BCD counters are conspicuous by their absence as they have generally been replaced by seven segment decoded counters. One disadvantage is a need in many cases for external drivers for LED displays but this will be eliminated when liquid crystal technology is more advanced and, hopefully, cheaper.

BINARY COUNTERS

As usual we will start with the less glamorous devices in the range which, in the present instance, are the straight-forward binary counters. First we should mention the general operating conditions required for all CMOS counters. The clock input rise and fall times should be less than 5 μs and the operating frequency limit is about 2.5 MHz at $V_{dd} = 5$ V rising to 5

MHz at 10 V. As far as the problem of drive current is concerned, it is advisable to consult the full data sheets for the device in question but it is reasonable to assume that no trouble is likely to be experienced if the requirement is less than 0.25 mA with a 5 V supply or 0.5 mA with 10 V.

Figure 1 gives the pin diagrams for CMOS seven, twelve and fourteen stage binary counters. The outputs are labelled B, with B_0 the most significant bit (i.e. giving greatest frequency division). It will be noted that three of the less significant bits are not available as outputs on the 4020A and this limits its usefullness in "divide by N" applications as we shall see later. The greatest division of the input frequency is 128 for the 4020A, 4096 for the 4040A and 16384 for the 4020A. In all cases the counters step on the negative transition of the clock pulse and the reset input sends all stages to logical zero independently of the clock when it is taken high. There is also a twenty-one stage counter (the 4045A) which produces two out-of-phase pulses at

separate outputs for every 2097152 input pulses. It is intended for producing one second pulses from 2.097 152 MHz crystals for driving clock circuitry and similar applications.

While we are on the subject of major frequency division chains perhaps we should consider crystal oscillators very briefly. Fig. 2(a) shows one common set-up and it is worth noting that the configuration in Fig. 2(b) is the standard way of producing a simple analogue amplifier from a CMOS inverter.

DIVIDE BY N COUNTERS

There are times when it is required to divide a signal by other than some power of two and by using a 4024A or 4040A we may divided by any number from two to 128 and 4096 respectively, although extra components are required. Figure 3 shows two ways of achieving this end.

The circuit in (a) has the binary counter feeding a system of logic gates, the output of which goes high when the counter reaches N-1 (where N is the number the input frequency is to be divided by). This happens on the falling edge of the clock pulse because the counters are negative-edge triggered. On the next rising edge the flip-flop Q output goes low and when the clock goes low again the output goes high, generating a pulse of length equal to one half of the clock period which resets the counter. It is interesting to draw a timing diagram for this circuit and prove it works. It should be noted that although the actual output is a positive going pulse, a similar pulse of twice its length (i.e. one clock period) is available at the Q output of the 4013. A divide by 3600 counter which will provide one pulse an hour from a 1 Hz input is
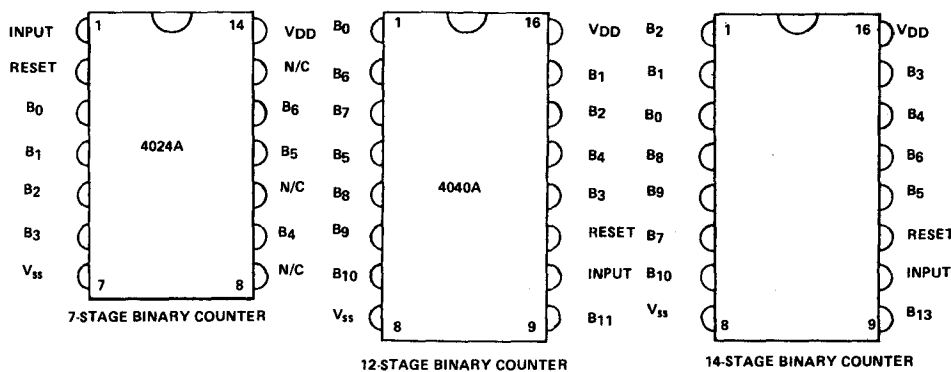
Fig. 1. Three CMOS binary counters.
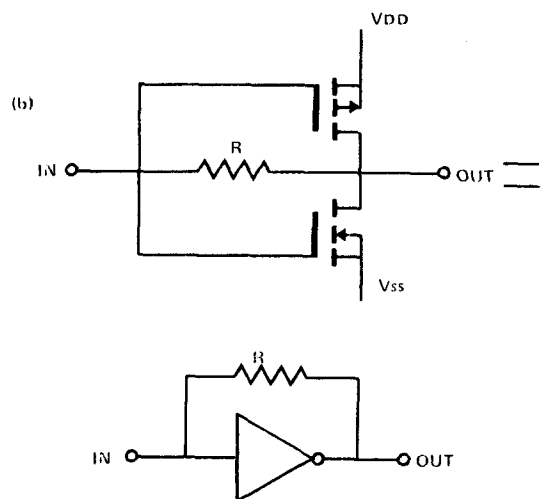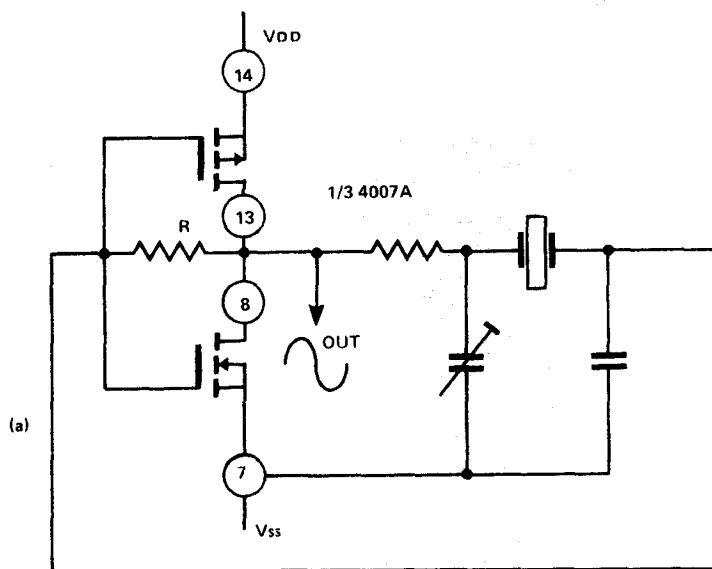
Fig. 2. (a) Basic crystal oscillator using CMOS for the active components.
(b) Simple analogue amplifier using a CMOS inverter.

shown in Fig.4 as an example of the technique.

The second mode has the advantage that the "N" count and not the "N—1" count is detected, but two logic networks are required; one to decide when the counter has reached "N" and another to identify the "all zeroes" state and reset the output. It is also a disadvantage in some applications that the counter spends a brief period in the "N" state. It is again interesting to draw a timing diagram and it is worth noting the cross-coupled NOR gates used as an R-S flip-flop. As an example a divide by twenty four counter is shown in Fig. 5 to produce one pulse per day from the one per hour output of Fig. 4 The circuit dissipation of both the counters would be very low (less than 1 mW) at this low operating frequency and the only note of caution to be sounded is that the counter and flip-flop should not both be triggered from the same edge of the clock pulse (i.e. one should be positive and the other negative edge-triggered).

A DECIMAL-DECODED DECADE COUNTER

All the old hands at TTL will doubtless be familiar with the 7490 decade counter and 74141 decimal decoder driver. The 4017A combines the count and decode functions in a single package but has the disadvantage of low output drive capability. Buffering the outputs with 4049A inverters will raise the available output to about five or ten milliamps at supply
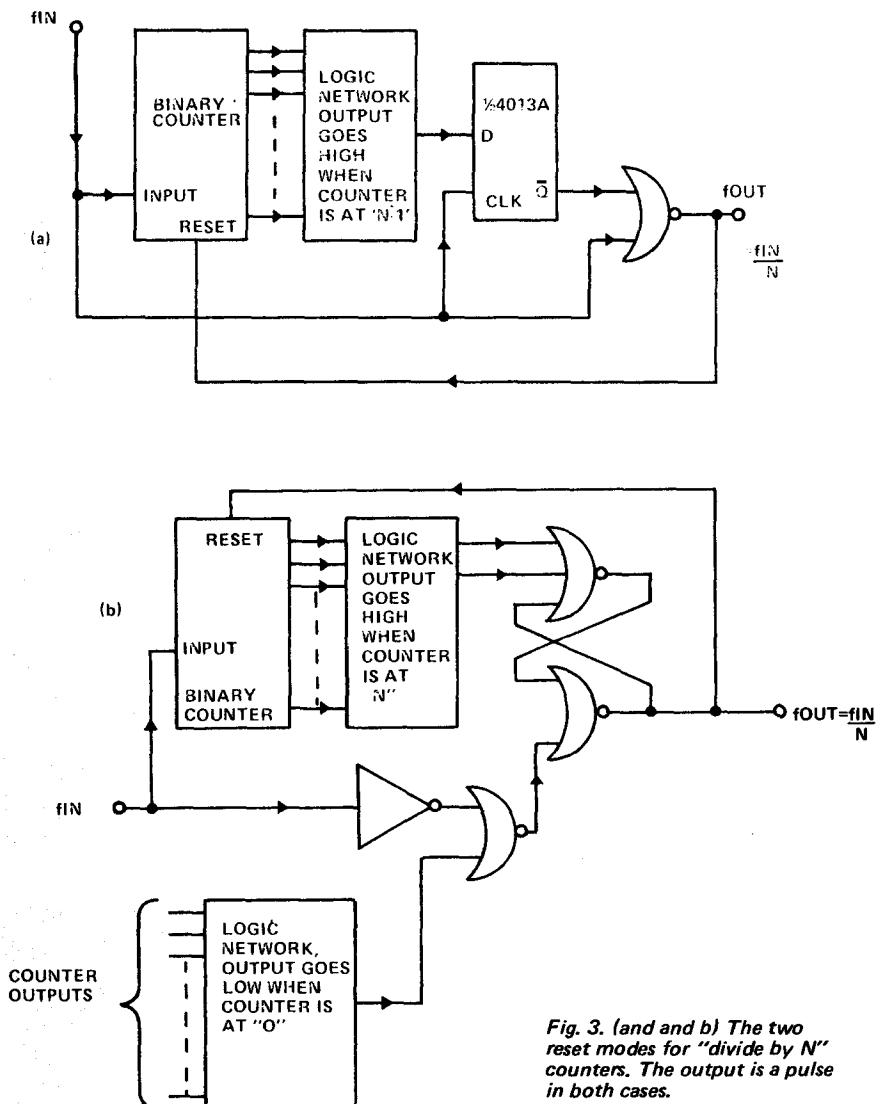
Fig. 3. (and and b) The two reset modes for "divide by N" counters. The output is a pulse in both cases.

CMOS-a practical guide
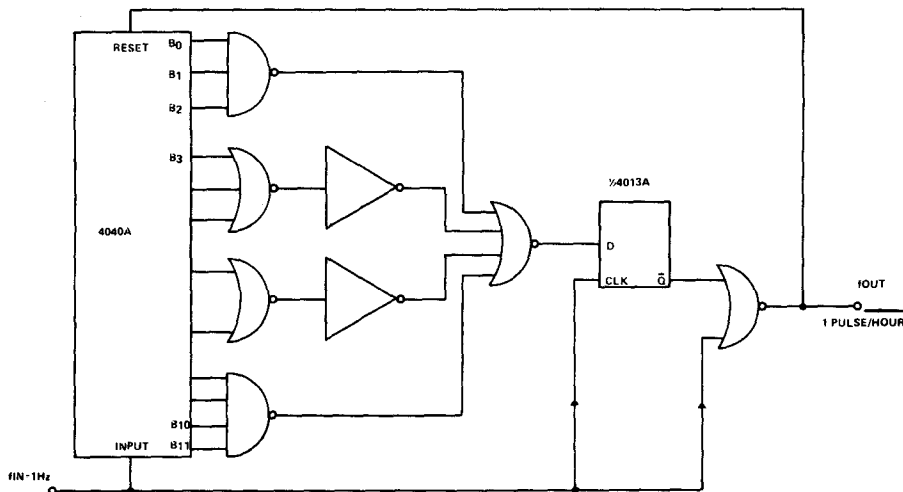
Fig. 4. A divide by 3600 counter using the first reset mode.
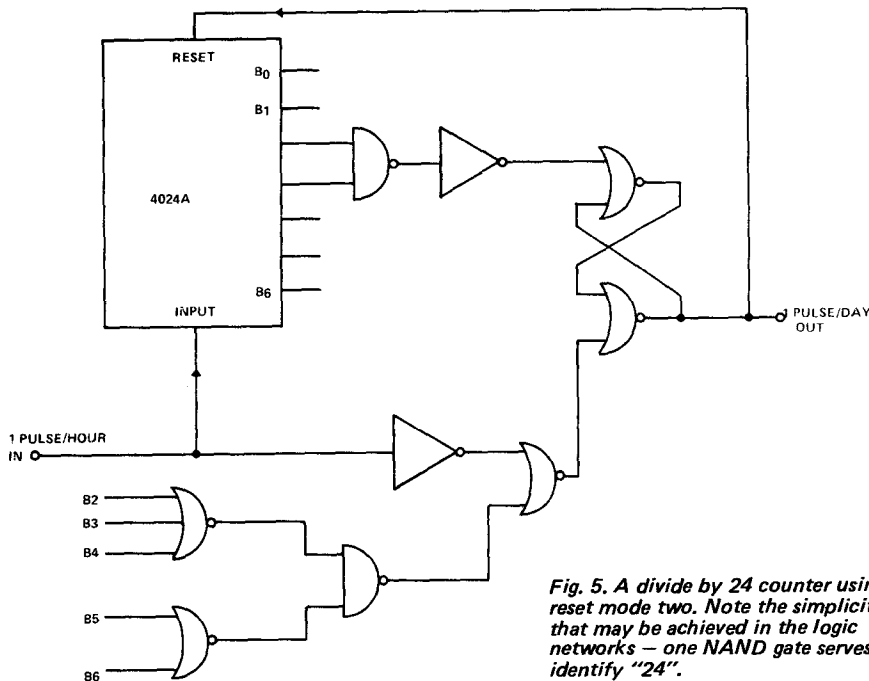
Fig. 5. A divide by 24 counter using reset mode two. Note the simplicity that may be achieved in the logic networks — one NAND gate serves to identify "24".

Fig. 6. Pin-out diagram of the 4017A decimally decoded decade counter.

voltages of five and ten volts respectively. The pin diagram is given in Fig. 6 and the counter advances one on the positive clock transition provided that the inhibit is held low. The reset operates asynchronously when taken high as usual. "Carry-out" may be used to clock the next stage in a multi-stage counter.

This device has fairly obvious applications in controlling switches in multiplying equipment as one and only one output is high at any one time. It is fairly clear also that we may extend the

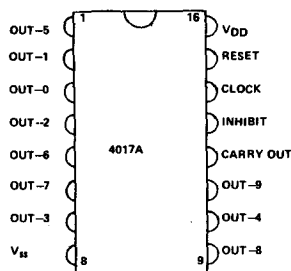techniques of divide by N counters to cover these devices with the added bonus that they are switch programmable. Figure 7 shows this idea realised using reset mode two because of the ease of switching for N rather than N−1. This circuit has lost an inverter compared with Fig. 5b, this being the change necessary to adapt the circuit for counters and flip-flops which operate on the same clock transition. The sequence of counters could clearly be extended to any desired length and it is an interesting thought that seven of these counters (4017As) and the attendant gates could, when fed with a 1 Hz input generate pulses at any interval from two seconds to over three months! On a more practical note a most versatile digital frequency synthesiser would result if the circuit were used on a phase-locked loop configuration. Remember however that the output is a pulse and it would need squaring (one more flip-flop) before most phase comparators would accept it.

SEVEN SEGMENT DECODED COUNTERS

We mentioned earlier that CMOS IC design reflected the changes in display technology. Two particular examples of this phenomenon are the 4026A and 4033A decade counters with seven-segment outputs. The pin-out diagrams for these devices are shown in Fig. 8 and, as one might guess, the counters are identical, with the exception that the 4026A has a display enable function for use in multiplexing digits and an ungated C-segment output, whereas the 4033A has ripple blanking and a "lamp-test" facility. We shall consider the use of these special facilities when we have discussed the features common to both. The devices are positive edge triggered and advance only when the clock enable is low. The reset operates when taken high as usual and the segment outputs go high when they are active. Just as in the 4017A the signal at the "carry out" terminal may be used to clock the next stage in multi-decade applications.

In the same way as we have considered for other counters, the seven segment outputs may be identified by logic gates and the counters made to divide by any number. Figure 9 gives the information necessary and it should be noted that the "N−1 and flip-flop" method is used because the other method does not count through zero.

Now we will have to consider the

Fig. 7. A switch programmable divide by N counter for N = 2 to 99. Extension to higher N is obvious.

Fig. 8. Pin-out diagram for the 4026A and 4033A seven segment decoded decade counters. The labelling of the segments is also shown.

Fig. 9. (a) How to produce direct seven segment divide by N counters, (b) logic networks to identify each digit. The extension to a multi-decade version is simple.

TO IDENTIFY	FOR COUNT OF	REQUIRED NETWORK
1	2	a-seg → X
2	3	c-seg → X
3	4	c-seg, g-seg → X
4	5	f-seg, g-seg → X
5	6	CARRY OUT → X
6	7	e-seg, f-seg, g-seg → X
7	8	g-seg, CARRY OUT → X
8	9	b-seg, e-seg, f-seg, g-seg → X

CMOS- a practical guide

Fig. 10. Driving MAN-3 type displays (a) at intermediate
supply voltages and, (b) at low voltage.

interfacing of displays with our seven-segment counters. LEDs like the MAN-3 which have a low current will interface directly with the outputs of the 4026A or 4033A and give a tolerable brightness with the available drive current (about 5 mA), provided that V_{dd} is more than 9 V. If we drop the voltage down to between 4 and 9 V transistors should be inserted, as shown in Fig. 10a, and if the supply drops even lower, the addition of inverting buffers is recommended. The seven transistors needed are generally the components of a single IC. Note also, the discussion on current limiting resistors to follow.

CMOS-a practical guide

Inherently rugged, CMOS logic has many advantages over other logic families — high noise immunity and uncritical power requirements are but two. This, the fourth article in this series, deals with multiplexing.
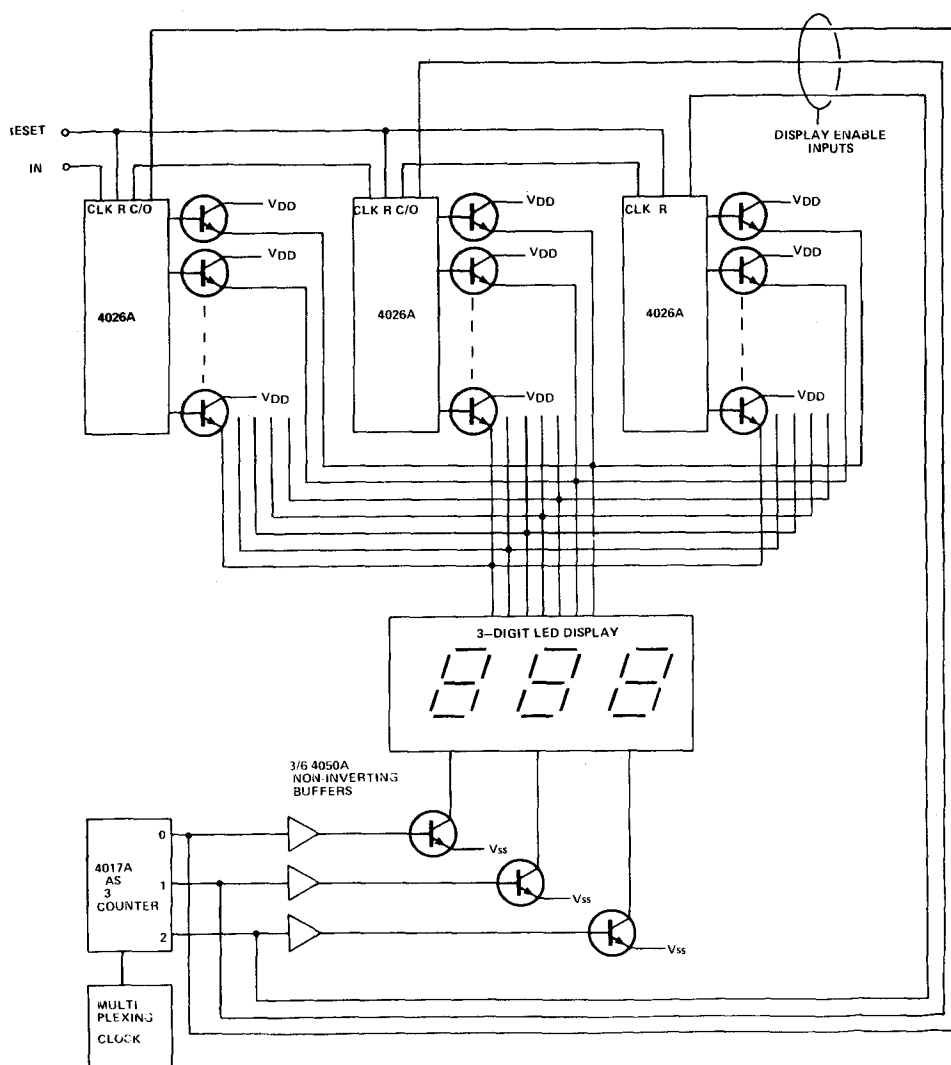
Fig. 1. A three decade counter for a 3-digit multiplexed display. Extra buffering of the digit lines may be necessary for some displays.

LIFE IS NEVER AS SIMPLE AS WE might want and there are two reasons for complicating the circuitry by using digit multiplexing (i.e. each digit is displayed for a fixed period, usually between about 10 and 30% of the time). These are that to do so is more efficient in terms of power consumption, and secondly that most multi-digit displays reduce the number of lead-outs (by giving just one set of seven segment drive lines for the complete display and one digit drive line for each digit).

This is the reason why the 4026A has a display enable input which, although the counter continues to function, cuts off the display when it is held low. The display enable output gives a replica of the input and may be used to enable other counters which are to be "on" during the same period. It also explains the presence of the "ungated C-segment" output which is used for producing some divide by "N" configurations which operate when the display is disabled.

The basic arrangement of a three decade counter is shown in Fig.1 and attention is drawn to the note that additional buffering may be necessary on the digit lines. It is also worth noting the use of a 4017 divide by three counter (using the flip-flop reset mode) to control the display.

Other types of displays which are often used are higher current LEDs such as the MAN-1 which is, in contrast to the MAN-3 a common anode device. This means it must be driven by inverting buffers as shown in Fig. 2a We have been relying here on the output current limit of the CMOS chip to limit
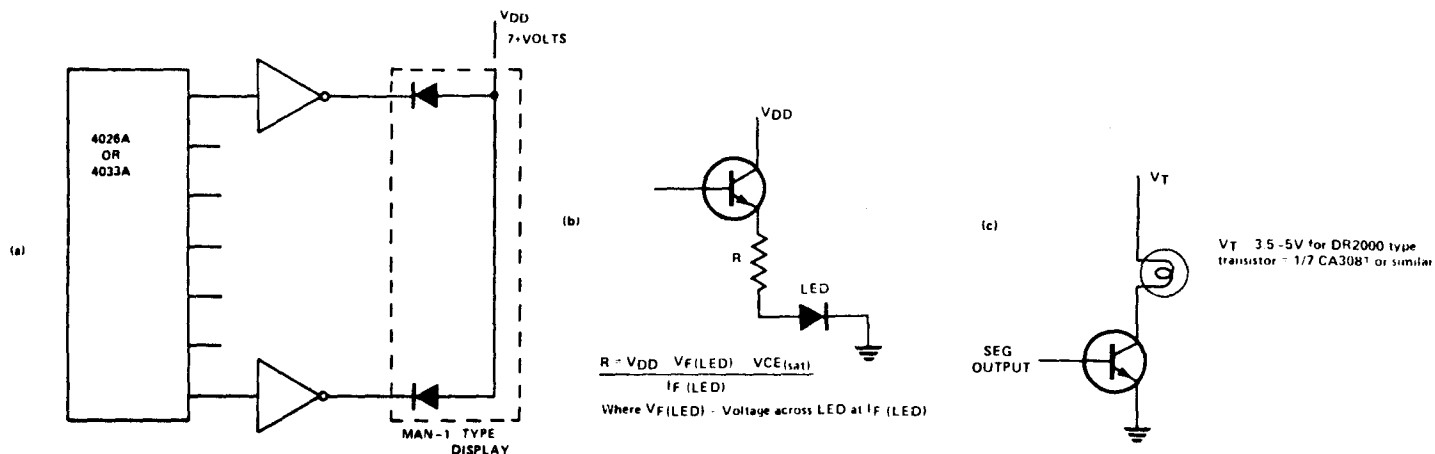
Fig. 2. Driving other displays: (a) MAN-1 type (b) example of calculation of limiting resistor (c) Numitron type incandescent display.

the forward current in the LEDs. It may be necessary to add current limiting resistors in the segment lines, particularly when transistor drivers are employed. The calculation of the value is simple given the required segment current and voltage drops (see Fig. 2(b). In multiplexed displays the limiting resistors should, of course, be put in the common segment lines. A considerable saving in resistors in non-multiplexed displays may be achieved by putting a single resistor in the common line to each digit. The disadvantage is that the display brightness varies with the digit. Fig. 2(c) shows the technique for interfacing with "Numitron" and similar displays.

The ripple blanking facility is for blanking leading and trailing zeroes in the display and it works as follows. Take the ripple blanking input (RBI) of the most significant 4033A on the integer side of the display low. Then take the ripple blanking output (RBO) of the IC and connect it to the RBI of the next counter and so on until the position of the assumed decimal point is reached. Follow exactly the same procedure from the least significant counter in the fractional part of the display backwards to the decimal point (see Fig. 3(a). Of course, if the assumed decimal point is at one end of the display then half the procedure would be unnecessary. If non-significant zeroes in the places either side of the decimal point are to be displayed (so that 7 and .6 appear as 7.0 and 0.6)

then the RBIs of the two counters concerned should be taken to Vdd. Finally on these two ICs, the lamp-test facility on the 4033A just forces all segment outputs high when it is taken high.

THE 4029A and 4081A

We shall conclude our discussion of counters by looking briefly at two more devices. The 4029A is a general purpose counter which, at the price that a 7490 was a year or two ago, has most of the features of the more exotic TTL devices. Briefly, the device is positive edge triggered and advances when the clock and preset enables are both low. Furthermore it counts in *binary* when the binary/decade input is high and BCD otherwise. A high signal at the up/down input persuades it to count up and a low input forces it to count down. As though this were not enough, when the preset enable input is high, the Q counter outputs are forced to follow the J ("Jam") inputs. The prefix "4" in both cases indicates the most significant digit. The pinout diagram is given in Fig.
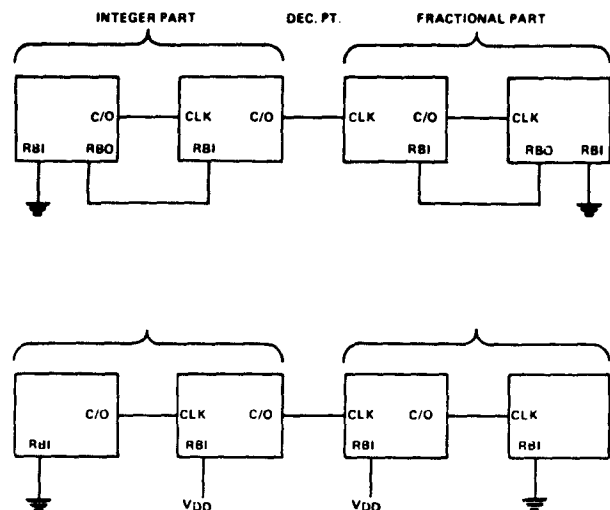
Fig. 3. Four digit counters using the 4033A with non-significant zero supression (a) in all positions (b) in first and last position only.
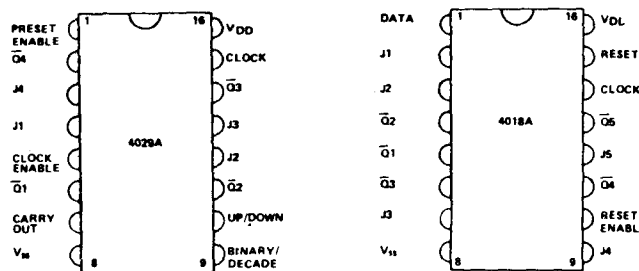
Fig. 4. Pinouts of the 4029A and 4018A.

CMOS-a practical guide
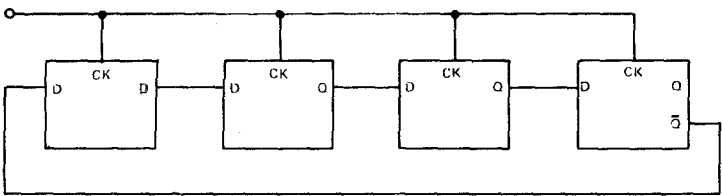
Q_1	Q_2	Q_3	Q_4
0	0	0	0
1	0	0	0
1	1	0	0
1	1	1	0
1	1	1	1
0	1	1	1
0	0	1	1
0	0	0	1

Fig. 5. Circuit diagram and counting sequence for a four stage Johnson or "twisted ring" counter.
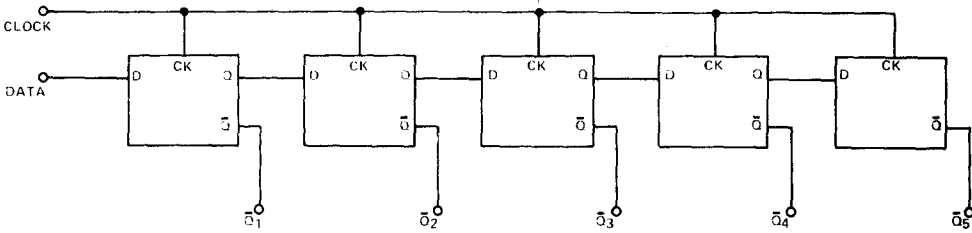
Fig. 6. Simplified internal diagram of the 4018A.

TO DIVIDE BY	CONNECT	
3	$\bar{Q}_1$ $\bar{Q}_2$	— DATA
4	$\bar{Q}_2$	DATA
5	$\bar{Q}_2$ $\bar{Q}_3$	— DATA
6	$\bar{Q}_3$	DATA
7	$\bar{Q}_3$ $\bar{Q}_4$	— DATA
8	$\bar{Q}_4$	DATA
9	$\bar{Q}_4$ $\bar{Q}_5$	— DATA
10	$\bar{Q}_5$	DATA

Fig. 7. Connection of the 4018A as a divide by "N" counter Input to clock, output waveform from DATA input is symmetric when N is even, almost so when N is odd.

4 along with that for the 4018A pre-settable divide by N counter.

There are two basic ways of producing counters. Firstly there is the chain of flip-flops each of which halves the frequency produced by the one before it. This was the principle behind the binary counters, which we considered at the beginning of this month's discussion, and also of the 4029A.

The second method is known as a Johnson counter and it is basically a shift register consisting of a chain of flip-flops with the Q output of the last counter connected back to the data input. A little patience and a pencil and paper will soon show that such a counter will divide the input frequency by 2N where N is the number of stages.

The counting sequence for a four stage counter is shown in Fig. 5 and the reader will notice that if the counter starts with contents not in the counting sequence (e.g. 1010) then the contents are always nonstandard thereafter. Thus some special gating is required. The simplified internal diagram of the 4018A in Fig. 6 is not complete. Also the Jam inputs and preset enable (which work in the same way as in the 4029A) together with the reset (which zeros all stages (Q1 — Q5 = 1) have been omitted for clarity.

Figure 7 shows the way to connect the 4018A to divide by all numbers from three to ten. Just as an example of how versatile this device is, one application will be considered in a totally different field from counting. By disregarding the clock, the Jam inputs and inverted data outputs ($\bar{Q}$) can be used as a five data latch for temporary storage, the outputs being updated to the inputs while the present enable is high.

CMOS – a practical guide

Shift registers, memories and the phase-locked loop.

WE HAVE ALREADY SEEN HOW flip-flops may be cascaded to form shift registers and, as you might guess, these are available already connected in a single IC. As usual we will start with some pin diagrams.

Referring to Fig. 1 we start with the 4015A, a dual four stage serial input parallel output (SIPO) version. The two parts "A" and "B" are independent and each consists of four D-type flip-flops with outputs Q1-Q4 and data is transferred from each one to the next and from the data inputs to the first on the rising edge of the clock pulse. A high input to the reset clears all the outputs to zero.

The 4006A is an 18 stage device with the registers so arranged that 4,5,8,9,10,12,13,14,16,17 and 18 stage registers may be produced. These are serial in-serial out (SISO) devices. It should be noted that the clock input is common to all the stage so that

although two separate five stage registers could be produced they would have to share the same clock.

Finally in Fig. 1 we have the 4014A and 4021A which are both eight bit registers with the same pin connections. The 4014A works as follows: On the rising edge of the clock pulse, the data is all shifted down and a new datum accepted from the serial input if the parallel/serial input is low. If instead it is high, no shift occurs and instead each stage of the register adopts the state present at the parallel inputs.

The 4021A is similar in almost all respects except that when the parallel/serial is taken high the stages are jammed asynchronously (i.e. it does not wait for the clock transition like the 4014A). We mentioned earlier that the 4042A quad D latch is often used for temporary storage but for many applications a shift register is superior,

giving options of serial, or parallel operation.

As far as eight bit data manipulation is concerned, the ultimate shift register must surely be the 4034A whose pinout is given in Fig. 2. It is described as a bi-directional bus register because the two data lines "A" and "B" are interchangeable in that if the A/B input is high the A lines are inputs and the B lines are outputs, whereas when A/B is held low the reverse applies. When the parallel/serial input is high then data is accepted from the parallel inputs (on the positive clock transition when asynch/synch is low, asynchronously otherwise) whereas when parallel/serial is low, data is accepted from the serial input and shifting is performed on the positive clock transition. The only other feature is the A-enable input which enables the A lines when it is high thus allowing several registers to feed a common bus. Designers of large digital

Fig. 1. Pinouts of four small shift registers.

A CLOCK — 9
A DATA — 7   A   5 — Q1
            4 — Q2
            3 — Q3
A RESET — 6
           10 — Q4

B CLOCK — 1
B DATA — 15   B   13 — Q1
              12 — Q2
              11 — Q3
B RESET — 14
               2 — Q4

V_DD 15. V_SS 8
4015A dual 4 stage SIPO

(1) IN — 4 STAGE — OUT (13)

(4) IN — 4 STAGE — 1 STAGE — OUT (12), OUT (11)

(5) IN — 4 STAGE — OUT (10)

(6) IN — 4 STAGE — 1 STAGE — OUT (9), OUT (8)

(3) — CLOCK
(7) — V_SS
(14) — V_DD   4006A 18 stage register

4014A Synchronous P/SISO register
4027A asynchronous PISO/Synchronous SISO register
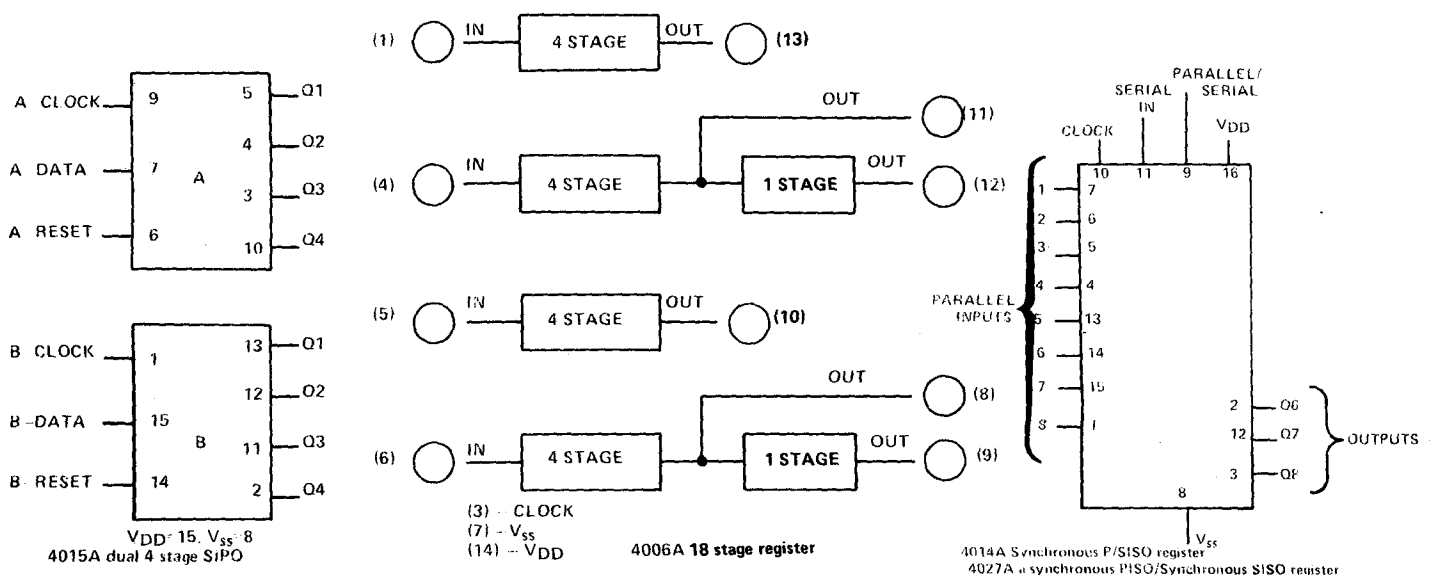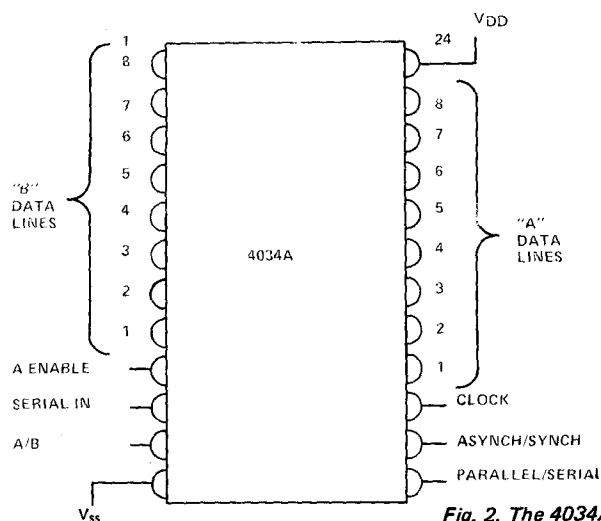
Fig. 2. The 4034A bidirectional bus register.

Q1	Q2	Q3	Q4
1	0	0	0
0	1	0	0
0	0	1	0
0	0	0	1

Fig. 3. (a) Johnson type waveform synthesis (b) principle of the single bit circulation register (c) counting sequence for (b) when it is correctly set up.
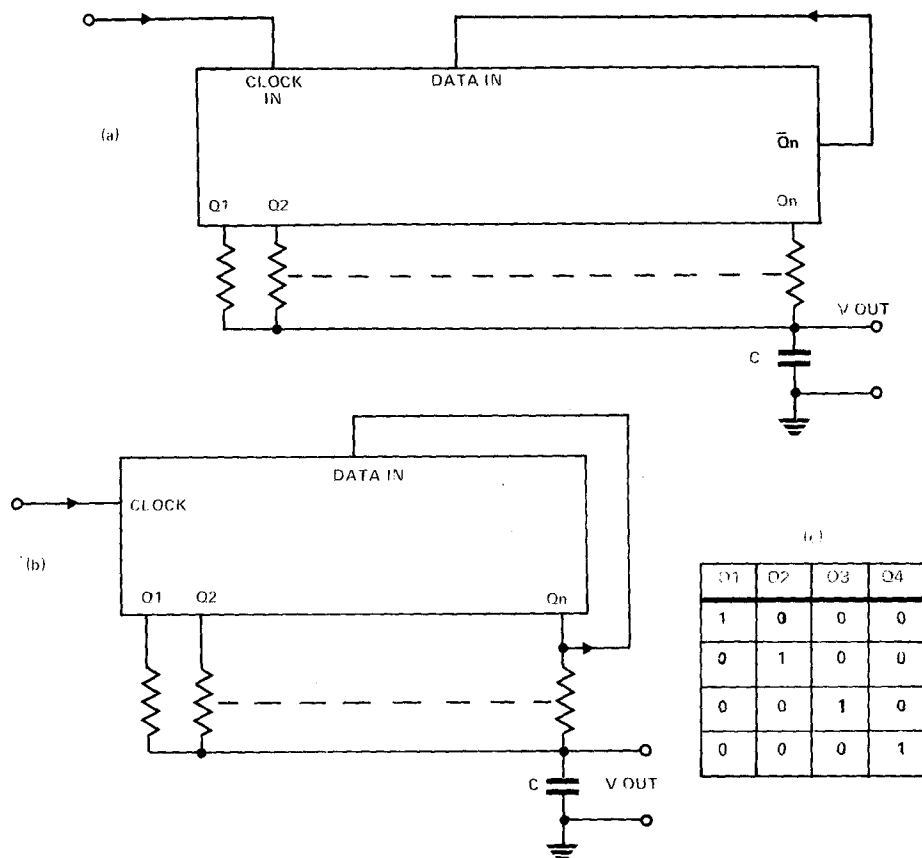
processing systems will find many applications of this device.

Applications of shift registers in data storage and conversion are not really very interesting out of context and so we will look briefly at counters. The connection of shift registers (under the guise of a sequence of flip-flops), to form Johnson counters, has already been considered but it is worth noting that in the same way as we produced digital to analogue converters from counters we can also produce waveform synthesisers from Johnson counters (see Fig 3(a)). The calculation of the required value of each register is rather
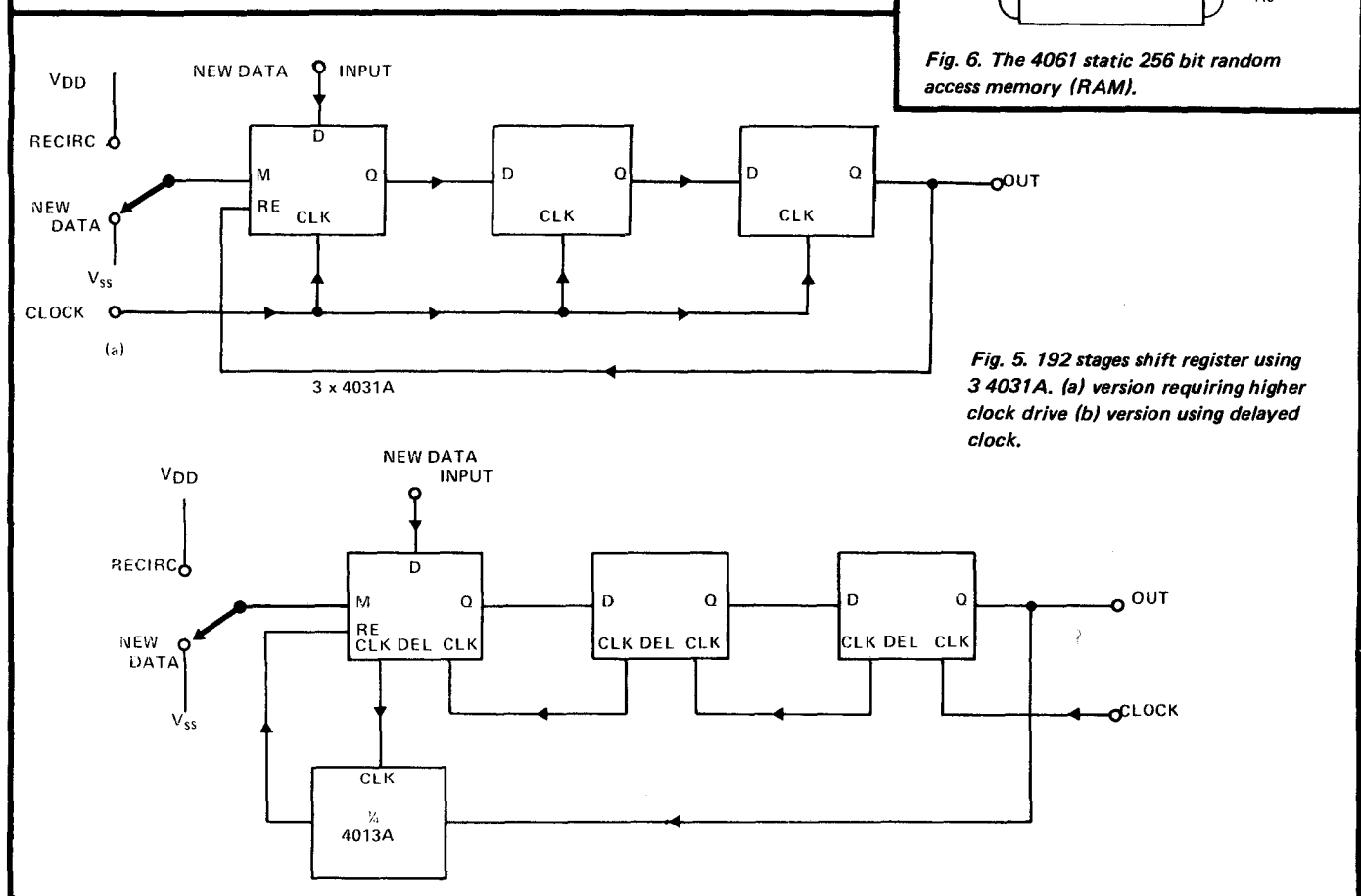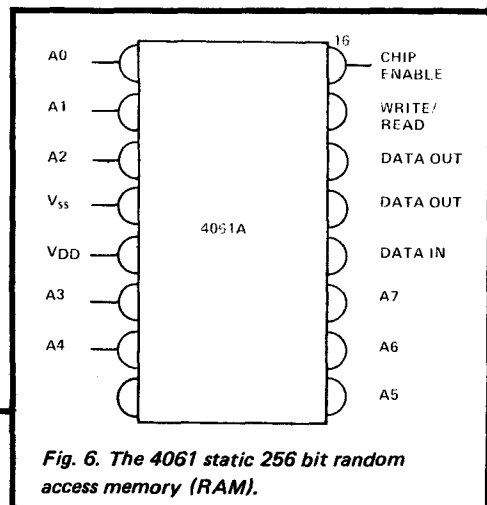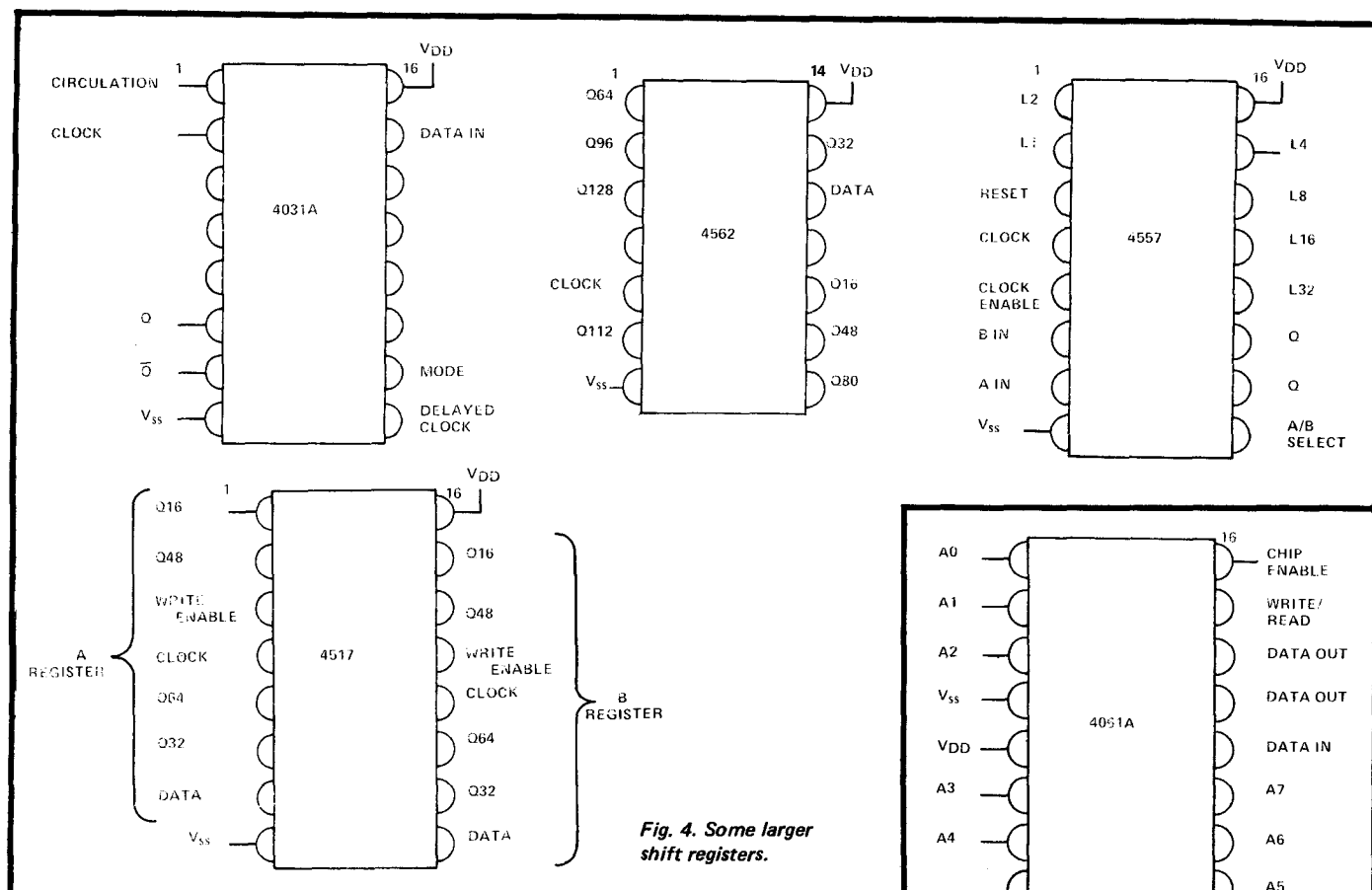
tedious because each is on for half a complete cycle. Nevertheless, an eight stage register gives sixteen separate values out but not all of these are independent. On occasions this method is superior to the single bit approach. The circuit in Fig. 3(b) is not complete because the single bit must be inserted to start with. If the application will tolerate switches in the starting sequence it is a trivial design exercise but if complete auto-starting is required it becomes quite interesting.

These small registers we have been considering are by no means the top end of the range. Figure 4 gives the pinouts for some high capacity registers.

The 4031A is a single sixty-four stage shift register, the only addition being a Q output and mode and recirculation connnections. When the mode control is low, the recirculation input is inoperative and data is accepted from the data input. When mode is high, however, data is accepted from the recirculation input instead. Thus if the Q output is connected to the recirculation input a register is formed with a control input which will either recirculate its contents continually or accept new data, depending on the control signal. You might well wonder why the recirculation input is not internally connected to the Q output. This is not done because it would generally reduce the versatility of the device and, more particularly, would prevent the formation of multi-chip recirculating registers like the one shown in Fig. 5(a). The 4013A is positive edge triggered and it has also got a delayed clock output which, as the name implies is a replica of the clock signal delayed by several hundred nano-seconds so that the cascading method shown in Fig. 5(b) may be used. The addition of the flip-flop if recirculation is required should be noted as without it this method will not work. The cascading method in Fig 5(a) must be used for highest speed operation although it requires higher clock drive.

Having said so much about the 4031A there is less to say about the next three devices. The 4517 is a dual sixty-four bit register, the two devices being distinguished in the pin diagram by the suffices "A" and "B". During normal operation the device is positively edge triggered, data is entered at the data pin and outputs are available at Q16, Q32, Q48 and Q64. This is with the write enable low. When it is high however data is entered at the data Q16, Q32 and Q48 pins (the Q16 pin acting as a

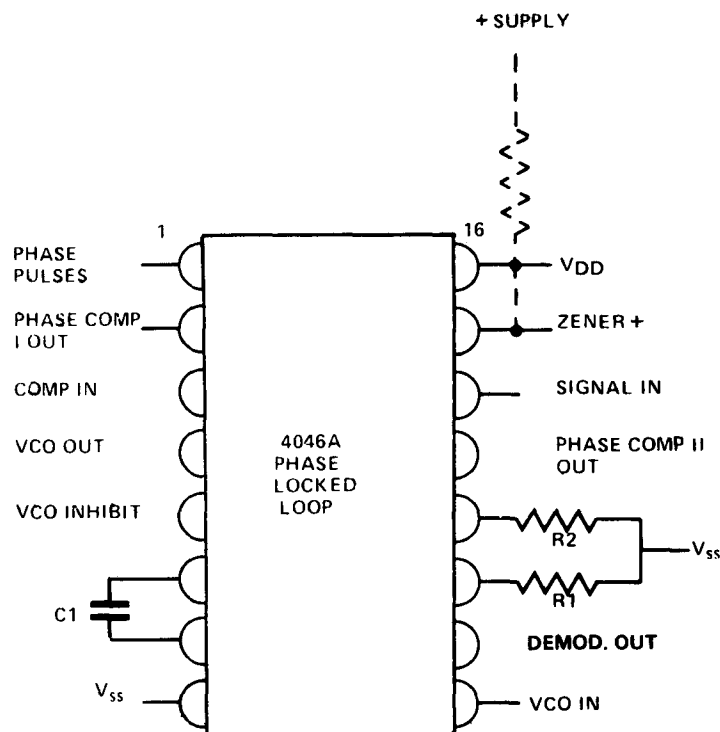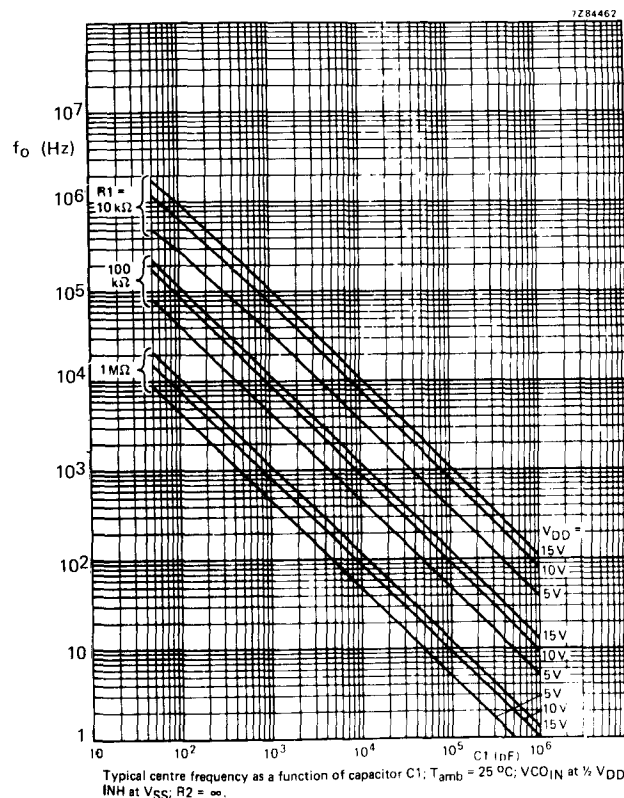# CMOS-a practical guide

Fig. 4. Some larger shift registers.

Fig. 6. The 4061 static 256 bit random access memory (RAM).

Fig. 5. 192 stages shift register using 3 4031A. (a) version requiring higher clock drive (b) version using delayed clock.

CMOS – a practical guide

Fig. 7. The 4046A pinout and chart for determining frequency.

Typical centre frequency as a function of capacitor C1; $T_{amb} = 25\,°C$; VCO_{IN} at ½ V_{DD}; INH at V_{SS}; R2 = ∞.

data input to stage 17 etc) thus allowing complete filling of the device in sixteen clock periods.

The 4562 is a straight forward one hundred and twenty eight stage positive edge triggered device with no additions except outputs every sixteen stages throughout.

The 4557 is an altogether more interesting device. Basically, the length of the shift register between the A and B inputs and the Q and Q outputs is equal to the sum of the L inputs which are high plus one. Thus if L32 and L4 are the only ones held high the shift register will be of length $32 + 4 + 1 = 37$. The reset input clears all stages to zero when it is taken high. The A/B select input is used to decide whether the A or B inputs will be used as the data source. When it is low B is used, when it is high A is chosen instead.

It should be clear from our previous discussion how a recirculating register could be formed. As for the clocking, when the clock enable is low the device behaves as a conventional positive edge triggered register, but if negative edge operation is required this can be achieved by holding the clock at logical one when a falling edge at the enable input will operate the device. Using this device as a Johnson counter would lead to an interesting binary programmable

symmetric frequency divider. Before we leave the subject of shift registers we will point out that all those we have considered are static devices which means that they maintain their state even if they are not clocked for some time. A two hundred stage dynamic shift register is available its type number is 4062A.

Of course, a shift register can be used as a memory. The disadvantage is that it is a serial access device which means that a lot of unwanted information must often be passed by while the piece we require is arriving. Random access memories (RAM) avoid this problem because access is by a set of address lines, the signal on which determines which bit of the memory is to be written to or read from.

A typical example is the 4061A two hundred and fifty six bit static RAM. This is simple enough for a brief description of its features to be given easily. A0 to A7 are the eight binary address lines necessary to specify any one bit and these should only be changed while the chip enable input is high. Only when the chip enable has returned to zero can read and write operations be performed — reading when "read/write" is low — writing otherwise. The buffering of the outputs is such that the data in line may be tied to either data out or data out if desired.

It should also be noted that the V_{SS} and V_{DD} are unconventionally positioned so that the devices are compatible with other MOS memories. Of course, there are a few other pieces of timing protocol to be observed. For example, the chip enable signal should not go low too soon after an address change.

THE 4046A PHASE LOCKED LOOP

The possibilities of combining linear and digital circuitry on the same chip has made the fabrication of the 4046A phase locked loop possible in CMOS. As usual we will discuss this device with reference to its pin diagram, which is given in Fig. 7.

A zener diode of 5.2 V is provided at pin 15 for supply regulation if required although the circuit will operate at any value of V_{DD} between five and 15 volts. The chip contains two phase comparators with outputs at pins 2 and 13. The first of these is an exclusive — OR network which requires a 1:1 mark space ratio at the input to work effectively, will lock onto harmonics of the input frequency but makes up for a lot by having good noise rejection and it also leaves the voltage controlled oscillator running in mid-range with no input connected. In contrast to this, phase comparator 2 can handle pulses
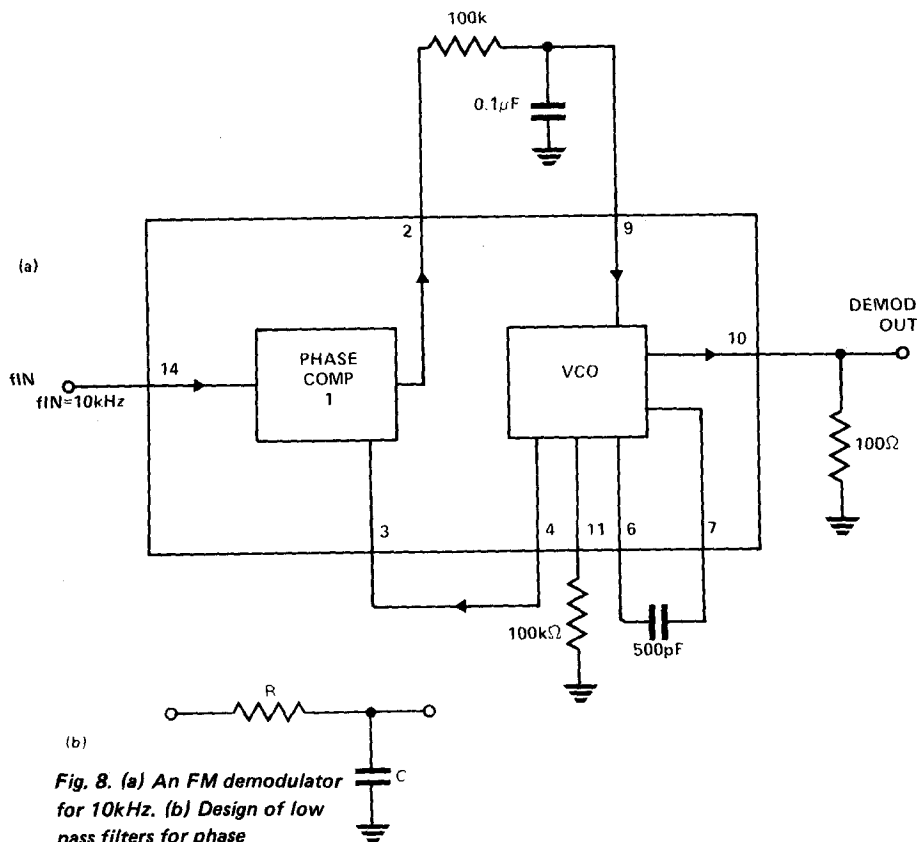
Fig. 8. (a) An FM demodulator for 10kHz. (b) Design of low pass filters for phase comparator 1.
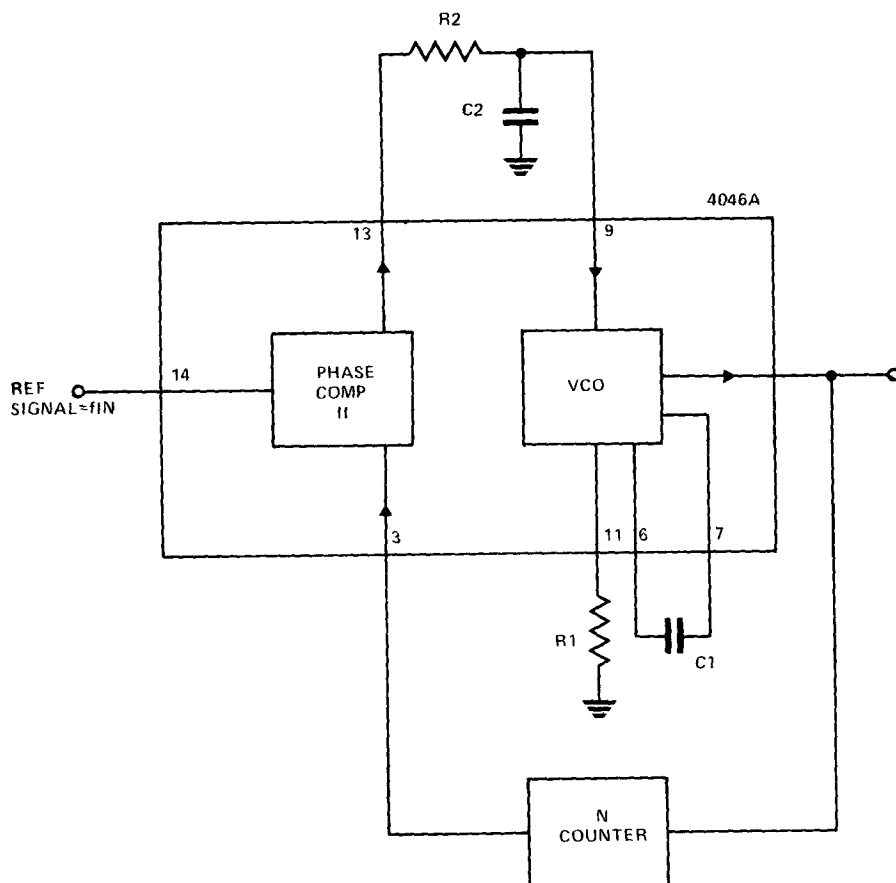

Fig. 9. Basic PLL frequency synthesiser.

and leaves the oscillator running at minimum frequency. The output from one of these phase comparators is generally connected via a low pass filter to the input of the V.C.O.

In passing we will mention that when the first phase comparator is used, the capture range (that is the range of frequencies that will pull the loop into lock) can be made less than the lock range (the range over which the loop will follow an already locked input) by using an appropriate low pass filter. In the case of comparator 2 however, the lock and capture ranges are identical.

The frequency of the VCO for given R1, R2 and C1 is shown in Fig. 7. The highest frequency attainable is over 1 MHz but the supply voltage may need to be a full 15 volts to achieve this with the cheapest plastic encapsulated devices. For experimenting purposes this description of the oscillator is probably sufficient but if it is required to design for specific frequency ranges, the R.C.A. data sheet on the device gives graphs which are easier to use than the somewhat cumbersome formula. Basically the smaller R2 is, the smaller the range of VCO may be swept over. If pin 12 is left open, the frequency can be swept from a negligible value to a megahertz or over.

If the signal in is within CMOS logic levels ("0"$\leqslant$ 30% V_{DD} "1" $\geqslant$ 70% V_{DD}) then it may be directly coupled to the signal input but if capacitive coupling is used signals at least as low as 1.5 V p-p will be accepted. The inhibit input saves power by disenabling the VCO when it is taken high and the "DEMOD OUT" terminal provides a buffered copy of the "VCO IN" signal for use as an output in demodulation set-ups. If this is used a resistor of 10 kΩ should be connected between it and V_{SS}. Figure 8 shows an FM demodulator designed to operate with a centre frequency of 10 kHz. It uses phase comparator 1 for good noise rejection and the formula for the capture range shows that the values used give ±400 Hz. As R2 is infinite the circuit will track over a very wide range and the linearity obtainable is very high.

Figure 9 shows the outline design of a frequency synthesiser. The second phase comparator is used because it will deal with non symmetric output from the divider and will not lock onto harmonics. The VCO is set with the pin 12 resistor infinite for maximum range. Three to nine hundred and ninety nine kilohertz steps can easily be achieved. The low pass filter should be optimised for minimum settling times. The divide by N counter has been shown as a block as these have been discussed at length previously.

CMOS-a practical guide

T Bailey completes his series on CMOS.

IN THIS FINAL ARTICLE WE WILL take a brief trip back through several of the topics already covered, looking at some more sophisticated ICs in each group.

The first two devices in Fig. 1 share a common pin-out diagram. They are both dual counters (labelled "A" and "B") with reset operating when high. The 4518 operates in BCD and the 4520 works in binary. Both devices are capable of counting at at least 2.5 MHz when V_{DD} = 10 V. The clock and enable inputs are interchangeable in that a positive edge triggered counter may be realised by holding enable at "1" and using the clock input, or a negative edge triggered device may be obtained by holding the clock high and using the enable input.

The 4553 is altogether a more advanced I.C. It is a three stage decade counter with latches and it provides a multiplexed output. The counters advance on the trailing edge of the clock pulse providing that "disable" is low. It will also advance on the rising edge of a disable pulse if the clock is high. The outputs are multiplexed, which means that one digit is given at a time on the

four BCD output lines. The three digit outputs show which digit is being presented (digit 1 is most significant). The BCD outputs are high when active, the digit-select outputs are low. The multiplexing is driven by an internal oscillator whose frequency is determined by the value of capacitor (1000 pF is about right) connected between pins 3 and 4. Alternatively, this can be overridden by leaving the capacitor out and driving the multiplexing by feeding pulses to pin 4. The carry out signal may be used to clock succeeding counters and in this case a capacitor may be used to control the multiplexing of the first counter and succeeding ones driven by connecting their pin 4 to pin 3 on the preceeding device. The reset input sets all the counters to zero and disables all the digit outputs hence blanking the display when it is taken high. The only other thing to note is the latch enable input. On the rising edge of the input to this pin the output from the counters is stored in latches and thus the conventional three decade counter ICs and three latch ICs are replaced by a single device. Use of this device is well

illustrated by the ETI counter module and also by Fig. 2 which shows a six decade version.

The two seven segment decoders used in these two counters – the 4543 and the 4511 – have their pin-out given in Fig. 3. The 4511 is a straightforward device with Q1-Q4 BCD inputs and a-g segment outputs. The three additional connections are simply a lamp test which lights all segments when it is taken low, a blanking input which turns off all segments, when it is taken low (unless lamp test is low as well) and a latch which stores the current input when it is taken high. The segment outputs will source up to 25 mA.

The 4543 is more advanced, the latch operates when taken low and the blanking operates when high. The device operates conventionally when the phase input is low (i.e. is suitable for directly driving common cathode LEDs) but when phase is high, the outputs are all inverted which is useful for driving common anode LEDs. If this input is fed with a square wave which is also fed to the common connection of the segments, liquid crystal displays may be driven in the manner described in part

Fig. 1. Three multistage CMOS counters.

"A" COUNTER

CLOCK
ENABLE
Q1
Q2
Q3
Q4
RESET
Vss

4518
4520

VDD
RESET
Q4
Q3
Q2
Q1
ENABLE
CLOCK

"B" COUNTER

Q1 is the least significant bit

DIGIT 2
DIGIT 1
MULTI-PLEXOR TIMING
Q4
Q3
Q2
Vss
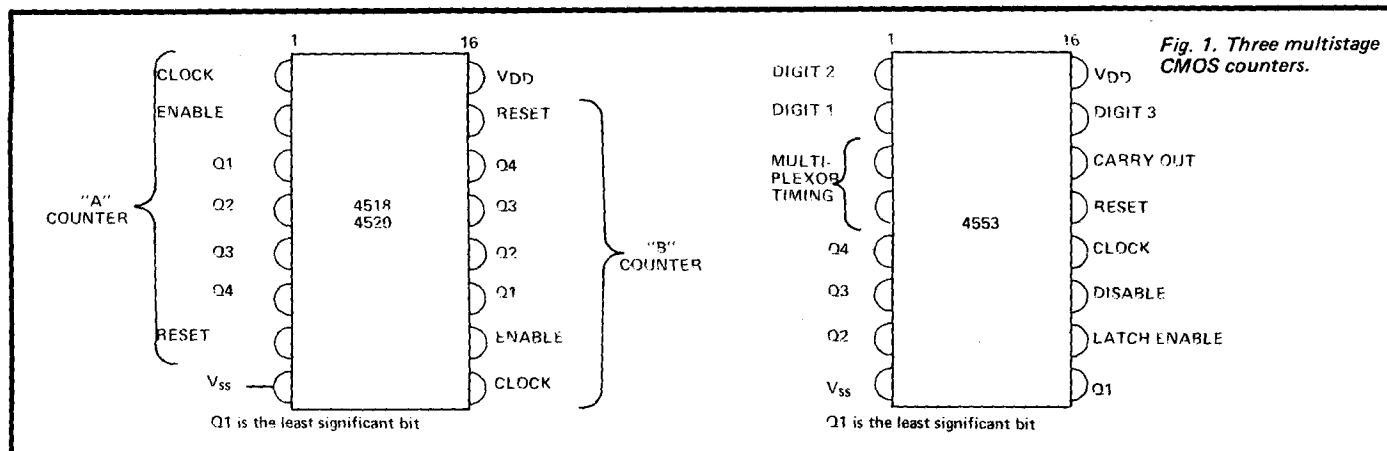
4553

VDD
DIGIT 3
CARRY OUT
RESET
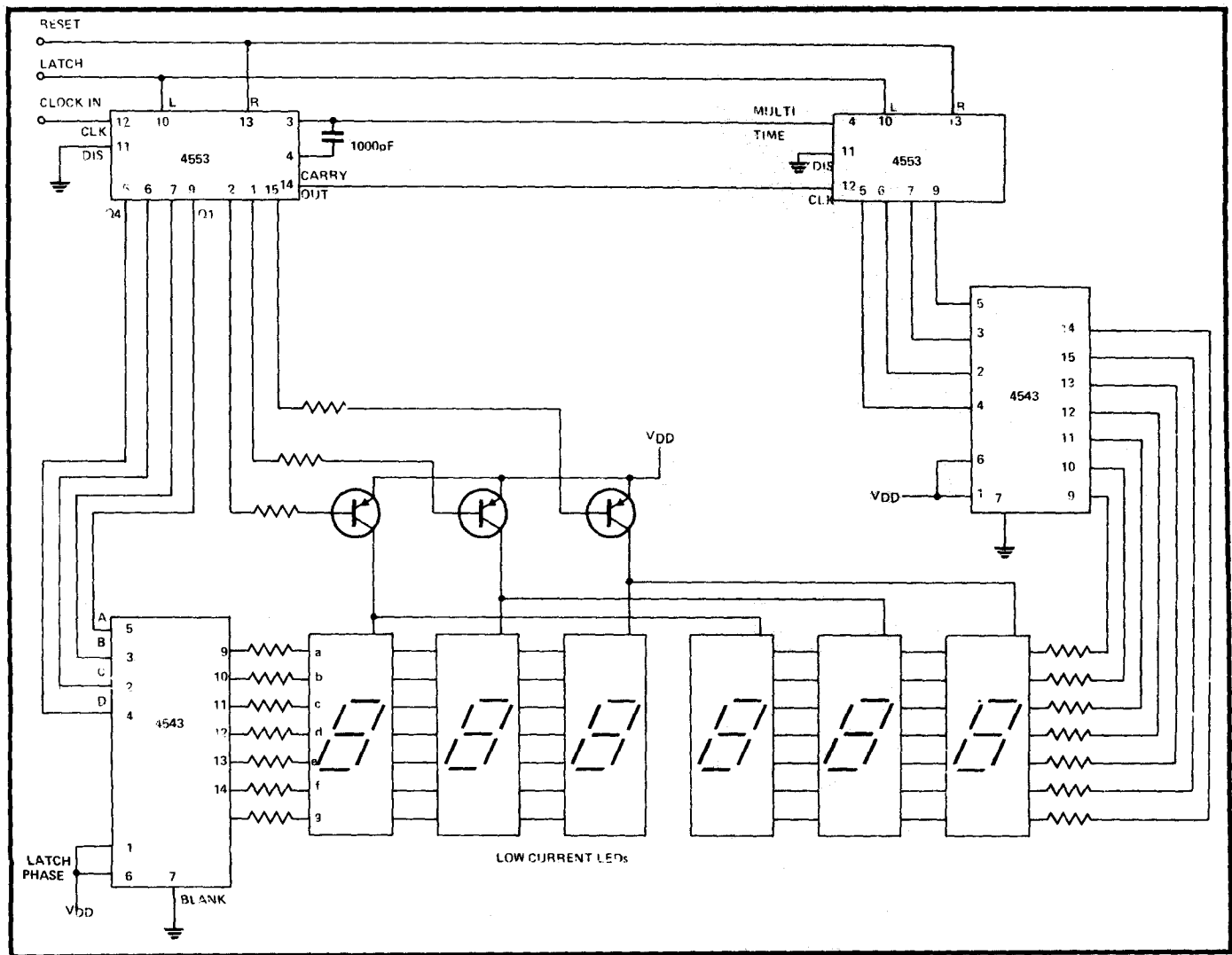CLOCK
DISABLE
LATCH ENABLE
Q1

Q1 is the least significant bit

Fig. 2. A six decade counter using CMOS. It may be adapted for common cathode LEDs by changing the drivers and taking pin 6 on the 4543 s to V_SS.

one. The 4056A mentioned there is a pin for pin equivalent of this device except that the blanking is dispensed with and pin seven used as a second V_{ss} pin for the display output part of the circuitry. Thus pin 16 could be at 0 V, pin 8 at -3 V and pin 7 at -15 V giving maximum economy while still providing full drive at the output.

There is also a five decade counter of a similar type but there is not space to describe it here. Its type number is 4534 and it comes in a twenty-four pin DIL case.
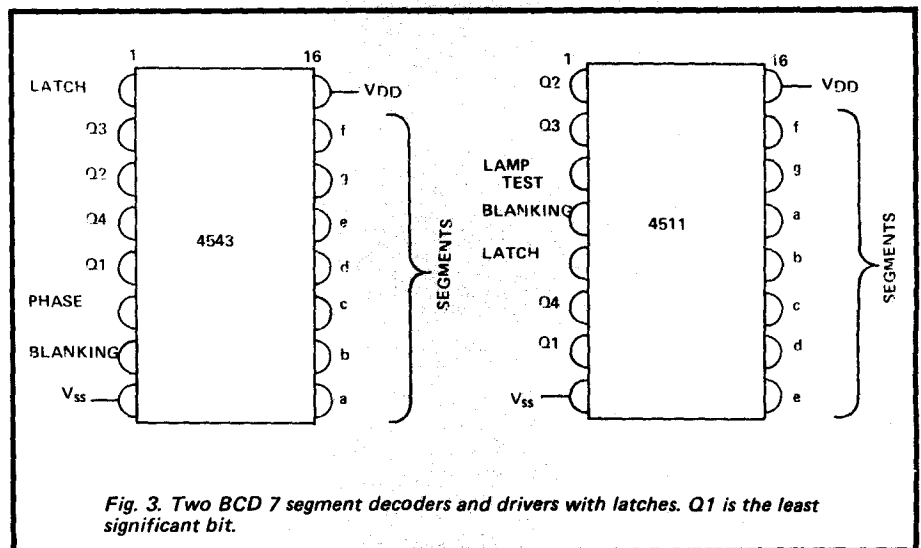

Fig. 3. Two BCD 7 segment decoders and drivers with latches. Q1 is the least significant bit.
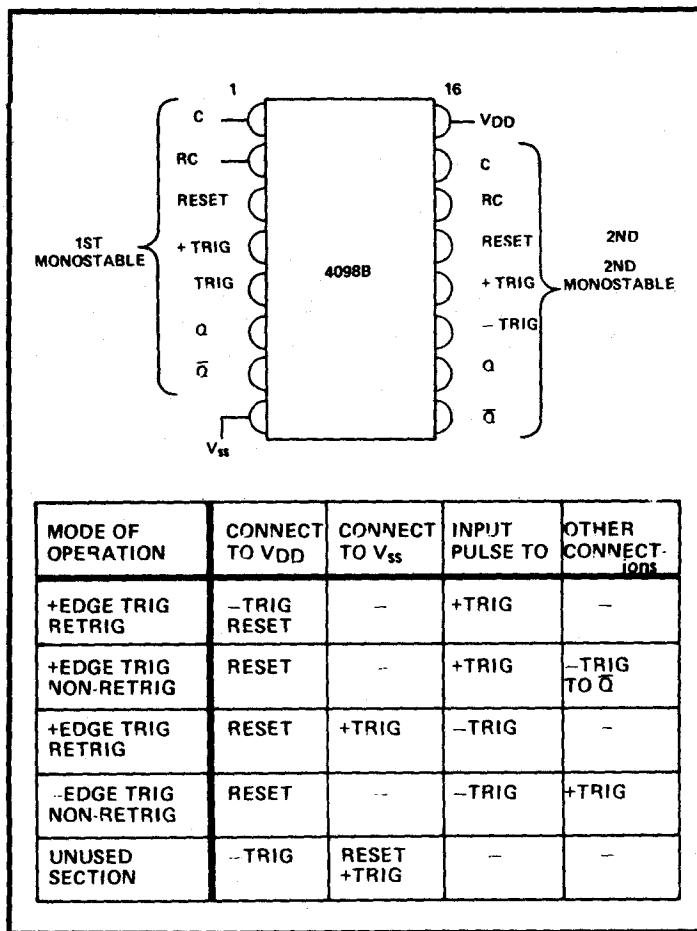
CMOS-a practical guide

MODE OF OPERATION	CONNECT TO VDD	CONNECT TO V$_{SS}$	INPUT PULSE TO	OTHER CONNECT-ions
+EDGE TRIG RETRIG	−TRIG RESET	–	+TRIG	–
+EDGE TRIG NON-RETRIG	RESET	–	+TRIG	−TRIG TO $\bar{Q}$
+EDGE TRIG RETRIG	RESET	+TRIG	−TRIG	–
−EDGE TRIG NON-RETRIG	RESET	–	−TRIG	+TRIG
UNUSED SECTION	−TRIG	RESET +TRIG	–	–

Fig. 4. The 4098B dual monostable multivibrator and method of achieving different modes of operation.

MONOSTABLE MULTIVIBRATOR

The 4098B is a dual monostable multivibrator. Its pin diagram (Fig. 4) is accompanied by a table showing the connections needed for every combination of edge triggering and retriggerability. The reset operates when low in this device whose period is, to a first approximation given by $T = RC$ (ohms and farads), where C is connected between the RC and C pins and R is connected from RC to V_{DD}. The specification of the 4528 is similar except for minor details.

MORE GATES

We can now claim to have covered a fair cross-section of devices and so to conclude we shall say a little more on the subject which we started with, simple gates. As well as the NOR and NAND gates we mentioned at the time there is a range of AND and OR gates available at comparable prices. The 4071B, 4075B, 4072B are quad, triple

Fig. 5. Pinout of the 4502 strobed hex inverting buffer.

and dual OR gates respectively with identical connections to the NOR gates (4001A, 4025A, 4002A) that were discussed in the first part of this series. Similarly, the 4081B 4085B, 4082B are the AND gates corresponding to the 4011A, 4023A, 4012A we mentioned then.

The 4030A quad exclusive-OR gate was also listed there and it is worth mentioning that types 4070B and 4077B are exclusive-OR and exclusive-NOR gates with identical pin connections. As the 4070B has slightly superior specification to the 4030A and is usually cheaper it may generally replace it. Also, for almost all purposes the 4507 is equivalent to the 4030A and 4070B.

The 4093B is a quad NAND Schmitt trigger with about 0.6 volts hysteresis (at $V_{DD} = 5$ V) and a pin-out identical to the 4001A. The 4583 is a dual Schmitt trigger in which the hysteresis may be adjusted by external resistors. There can be few uses for these which have not already been realised with the TTL SN7413N but it is worth noting that larger time constants could be used on the inputs.

Figure 5 shows a hex inverter and buffer with the extra options of an inhibit input which makes all the inverters have low outputs when it is taken high and an output disable which sets all the outputs in a high impedance state. This also operates when it is taken high. The chief use of these circuits is in applying one of two lines of data to an input. They are both wired in but only one disable is low at any one time. The disable overrides the inhibit.

RANGES OF CMOS

Throughout this series, devices have been known by a four digit code number beginning 40 and ending with A or B, or beginning 45 and possessing no suffix. Most of the devices beginning 40 are available from RCA in the CD range with a type number CD40xxAE or CD40xxBE. The A signifies that the maximum supply voltage is 15 V, B signifies 18 V. In general, A and B versions are not both provided. Most of this range is also available from Motorola as thMC140xxCP range which will tolerate up to sixteen volts. The 45 devices are often available only in the MC145xxCP range. In general other combinations of suffixes indicate a ceramic packages or the like. Generally these are more expensive and have slightly superior specifications.
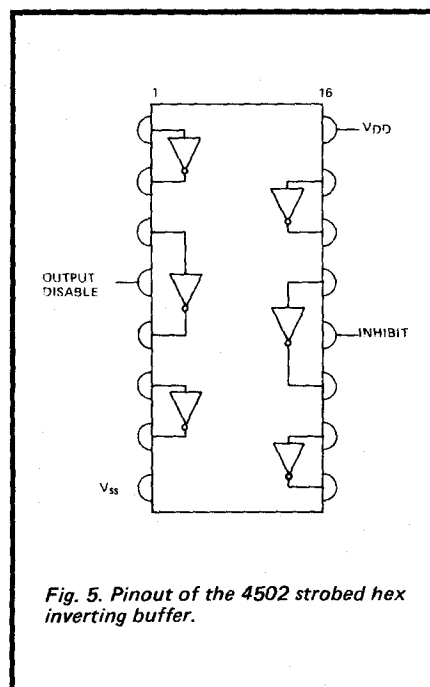
Lab Notes

Safety with CMOS

Certain elementary safety precautions must be taken when handling CMOS ICs or designing CMOS circuits. Ray Marston explains.

EARLY CMOS ICs earned a reputation for being easily damaged by static electricity, either when being handled or when being soldered into circuit boards, etc. Subsequently, manufacturers tried to overcome this 'fragility' problem by providing the ICs with extensive built-in input and output protection on each gate in each package. These protection networks do a fairly satisfactory job, but provide the designer with a few extra problems when employing CMOS circuits.

CMOS protection networks

CMOS ICs are, by definition, metal-oxide semiconductor devices, in which the input signal is applied to the near-infinite impedance (about 10^{12} ohms) of the metal-oxide gate. Typically, the gate oxide has a breakdown voltage of about 80 V; if a gate oxide breakdown does occur, the resultant damage to the device is catastrophic and irreversible. To protect the CMOS against excessive input voltages (particularly arising from static energy), all modern CMOS

ICs are provided with extensive built-in protection on all inputs and outputs.

Figure 1 shows the standard protection network that is used on the vast majority of B-series CMOS devices. Here, all diodes marked as 'D1' are used to prevent the input or output from swinging more than 600 mV below the V_{SS} (0 V) rail, and all diodes marked as 'D2' are used to prevent the input or output from swinging more than 600 mV above the V_{DD} (supply positive) rail. D3 is intended to prevent the V_{DD} terminal from swinging negative to the V_{SS} pin (electrostatically) when the device is being handled.

There are a couple of minor exceptions to the standard version of the protection network. One of these is the type used on the 4049B and 4050B series of hex buffer/converters which, as shown in Figure 2, have their inputs free to swing well above the V_{DD} rail. These particular ICs are specifically intended for use in logic-level conversion applications, in which (for example) the input may come from a 12 V CMOS net-

work but the output and the IC supply rail are matched to a 5 V TTL network.

Another exception is the 4066B type of transmission gate or bilateral switch, and its equivalents. These devices comprise a bilateral electronic switch and a switch-control network. In these circuits, all switch-control networks have the type of input protection shown in Figure 1, but the switches themselves have the simple protection network shown in Figure 3.

Note in Figures 1-3 that all diodes marked with asterisks are 'parasitic' devices, which just happen to occur fortuitously as an inherent part of the CMOS manufacturing process, while all other diodes are specifically designed into the circuits. Also note that the networks are intended only to give protection against 'normal' electrostatic discharge voltages. When the networks are subjected to ordinary dc signals, the diodes are liable to burn out if their forward currents exceed 10 mA or so, thereby causing possible catastrophic damage to the IC substrate. ▶
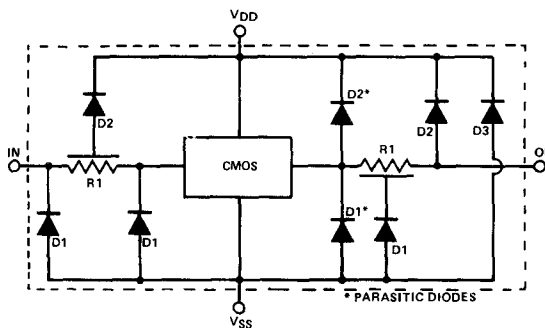

Figure 1. These are the standard electrostatic discharge protection networks used on most B-series CMOS ICs. The two diodes associated with the resistors are distributed across the entire resistance, as shown.
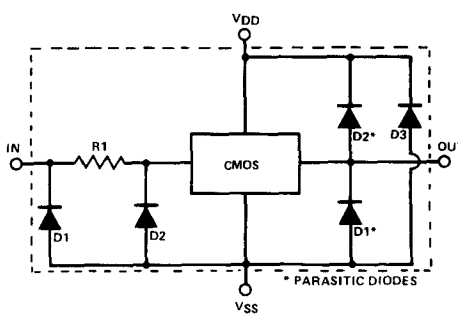
Figure 2. This protection network is used on the 4049B and 4050B hex buffers. Note that the input is free to swing above the positive supply (V_{DD}) rail.
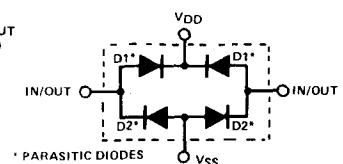
Figure 3. The 4066B quad bilateral switch has standard B-series protection on its gate control input terminals, but has this simplified form of protection on its 'switch' elements.

Lab Notes

Major CMOS manufacturers such as RCA reckon that an electrostatically charged human body can be approximated by the circuit of Figure 4, in which the 'body' has an effective capacitance of 100 pF and a source resistance of 560R. The manufacturers have carried out extensive tests with this model by charging the 'body' to various
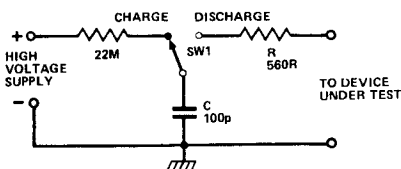

Figure 4. **Manufacturers use this equivalent-body discharge network when evaluating the capabilities of their CMOS protection networks.**

voltages and then discharging it (via the 560R series resistor) into different terminal combinations (input, output, V_{SS}, V_{DD}) of CMOS devices to establish worst-case capability figures for the three types of electrostatic-discharge protection networks. It should be noted in these tests that the 560R series resistor acts as a current-limiting voltage dropper, so the voltage actually reaching the CMOS device is far lower than the initial electrostatic voltage.

The results of the manufacturer's protection capability tests are shown in Figure 5. As you can see, the standard protection network can withstand a 4 kV electrostatic discharge. A quick calculation shows, however, that this represents a peak protection-diode current of several amps, yet we've already seen that these diodes can withstand dc currents of only 10 mA or so. Puzzled?

PROTECTION NETWORK	WORST – CASE CAPABILITY
STANDARD B–SERIES 4049B AND 4050B 4066B BILATERAL SWITCH	4 kV 1 kV TO 2 kV < 800 V

Figure 5. **These are the worst-case capabilities of the three different CMOS protection networks, when tested with the network of Figure 4.**

Up the junction

Just about the only way of destroying a diode is to literally vaporise its junction, and this can only be done by applying an

adequate amount of power for sufficient time for the melting process to take place. Since a junction must inevitably be formed on a substrate, which has a finite mass, all junctions inevitably have a certain amount of thermal inertia and are, in fact, destroyed by energy overloads (power-time product), rather than by simple power overloads.

Consequently, it is quite normal to find that a diode rated at 1 A (for example) can, in fact, withstand brief current surges up to several hundred amps. Similarly, CMOS protection diodes, which have very low dc current ratings (10 mA), can withstand very high levels of surge current (several amps), provided that the surge current
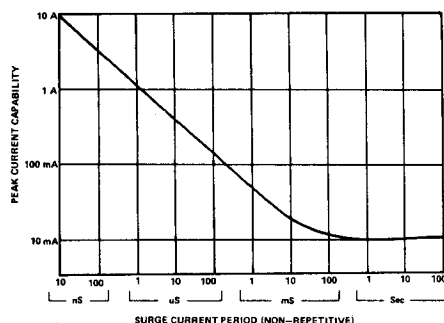

Figure 6. **Typical surge-current capabilities of CMOS protection diodes.**

duration is very brief. Figure 6 shows the typical surge current capabilities of these protection diodes. Remembering that the 100p — 560R 'human body' equivalent circuit has a time constant of a mere 56 nS, it no longer comes as a surprise to note that these diodes can withstand several amps of peak current from a 4 kV discharge!

CMOS circuit design

By now you will have gathered that you can effectively destroy a CMOS device by simply blowing one or more of its 'protection' diodes with a dc current as low as 10 mA. Consequently, when designing CMOS circuits, precautions must be taken to ensure that excessive diode current cannot flow in the CMOS chips.

CMOS ICs can be 'blown' by excessive signals applied to either the input or the output terminals. If several CMOS stages are cascaded, empirical ex-

perience shows that a front-end 'blow' will usually destroy only a single device (because low energy levels are normally involved), but a rear-end (output) 'blow' will often have a ripple effect (because high energy levels are involved) and cause the destruction of all ICs in the chain.
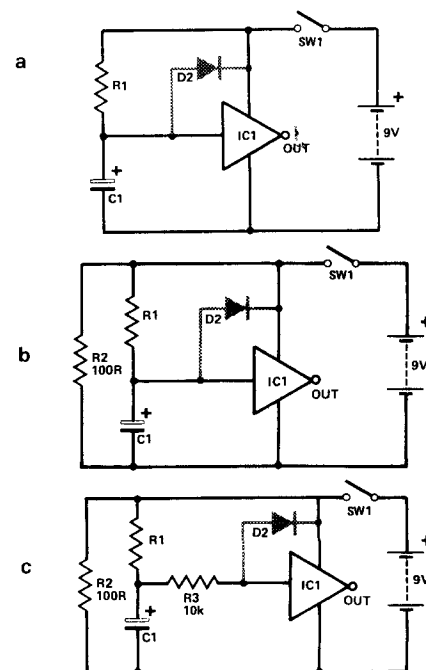

Figure 7. **Circuits (a) and (c) are safe, but circuit (b) will almost certainly cause front-end 'blow'. See text for explanation.**

The most common cause of front-end 'blow', and its cure, are illustrated in Figure 7. Here, a capacitor is connected directly between the IC gate and the 0 V line; when SW1 is closed, the capacitor charges up via R1 and eventually attains the full positive supply potential. When SW1 is opened (to switch the circuit off), C1 tries to discharge via D2, the 'upper' input protection diode of the gate.

In the Figure 7a circuit, the only discharge path for C1 is via D2 and the IC's supply terminals; consequently the discharge currents will be quite low and the IC will probably suffer no damage. In Figure 7b, on the other hand, a 100R resistor is connected across the supply terminals, so C1 will try to discharge to ground via D2 and R2, and the resulting 90 mA peak current will almost certain-

ly result in the destruction of the chip. In practice, R2 may well take the form of various resistors and semiconductor devices distributed throughout the total circuit.

Figure 7c shows the cure for the Figure 7b design problem, a 10k resistor wired in series with the gate to limit the C1 discharge currents to a safe value. Whenever you design CMOS circuits and have to connect a capacitor between a gate and the 0 V rail, always make sure that the capacitor discharge current is limited to a safe value, either by a series gate resistor or by some other factor.
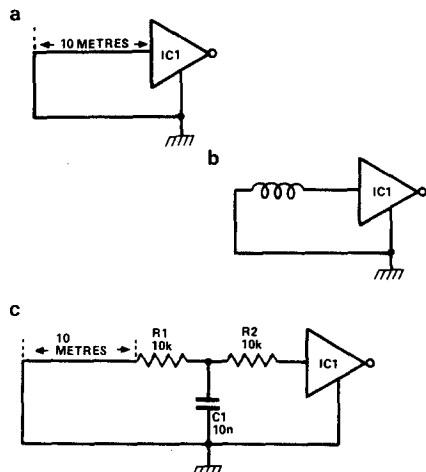
Figure 8. Long input cables, as in (a), can be equivalent to an inductor (b), and present another front-end blowing hazard. The cure is simple (c).

Figure 8 illustrates another possible cause of front-end 'blowing', and its cure. In Figure 8a, it seems that the IC's input is safely grounded by the 10 m of input cable (in practice, this cable may go to a low impedance sensor, etc), but in actual fact (Figure 8b) this cable will inevitably be inductive and can easily pick up unwanted radiation and possibly feed destructive signals to the IC input. Figure 8c shows that the circuit can be rendered safe with a simple filter (R1-C1) and a series gate resistor (R2).

Back-end blowing

The most common cause of back-end blowing is unexpected back-EMFs (from inductive loads) reaching the CMOS output by breaking through from power-driving circuitry.

Inductive loads, such as relays, can generate surprisingly large back EMFs as their fields collapse at switch-off, as can be proved by connecting a relay in the 'buzzer' mode shown in Figure 9. Typically, a 12 V relay will generate a back-EMF of about 300 V! If you ever use CMOS to switch a relay or other highly inductive load using a transistor driver, always protect the transistor with a pair of 1N4001 diodes connected as shown in Figure 10a. If you want to be really safe, you can use another pair of similarly connected diodes to directly protect the output of the CMOS stage, as shown in Figure 10b. ●
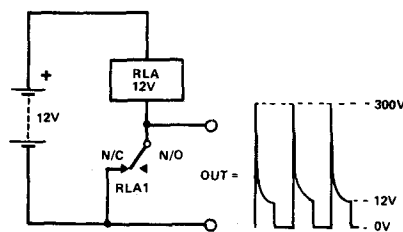
Figure 9. This 'buzzer' circuit can be used to check the magnitude of the back-EMF from a relay. 300 V is typical!
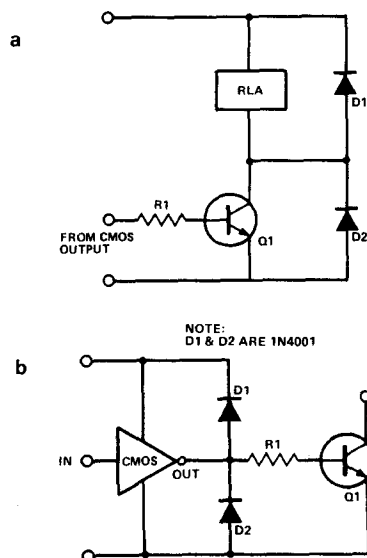
Figure 10(a). A transistor relay-driver can be protected with a pair of diodes. (b) The output of a CMOS stage can be given added protection with a similar arrangement.

ALL TOGETHER !

— the top projects from the past few years' issues of ETI !

Assembled in one great volume, here are more than 25 projects including: the Series 4000 stereo hi-fi amplifier system: the superb Series 4000 four-way speaker system: the monstrous 300 W amplifier module: Guitar Practice Amp: Fuzz/Sustain Unit: Electronic Tuning Fork; Aircraft Band Converter; Analogue Frequency Meter; Linear Scale Capacitance Meter; Electronic Ohmmeter; Ultrasonic Switch: Radio Remote Control Unit: Transistor Assisted Ignition: Theatrical Lighting Controller: Electromyogram: Disco Strobe . . . and many more.

On sale now at newsagents and selected specialist suppliers. If sold out (highly likely), you can obtain a copy from **Electronics Today International, Subscription Department, 3rd Floor, 15 Boundary St, Rushcutters Bay NSW 2011.** *Send $3.95 plus 65 cents post and handling.*

VFETS FOR EVERYONE

Wally Parsons looks back to valves to explain VFETs.

A DIODE VALVE emits electrons from a heated cathode and these are then attracted by an electric field to the positive anode. Since only the cathode is heated, current can flow in only one direction. The diode will thus act as a rectifier, conducting only on alternate half-cycles of an AC voltage (see Fig. 1).

If a grid structure is placed between these electrodes, it can be used to control current flow. A negative potential will repel electrons, opposing their flow to the anode, and by placing the grid close to the cathode, a small change in grid potential will have the same effect on anode current as a much larger change in anode potential. Therefore, the device will amplify. Since the anode current is controlled by the electric field in and around the grid, the triode is, in a sense, a field effect device.

The action is direct, and electron flow responds rapidly to changes in control potential. Moreover, in switching applications it can switch an inductive load rapidly, because the back EMF sees an extremely high impedance and no reverse current flows.

Figure 2 shows the relationship of anode voltage, grid volts, and anode current for a triode. It can be seen that anode current can be controlled by both anode volts and grid volts. If a load is inserted in the anode circuit, current changes will cause voltage changes across the load. These can be plotted in the form of a load line as shown, and also as a transfer curve for the specific load.

The amplification is quite linear, but gain and output are limited — as shown by the semi-vertical slope of the curves.

Inserting a second grid between the control grid and anode and applying a fixed positive voltage somewhat lower than on the anode further accelerates
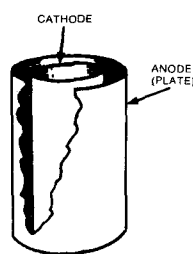

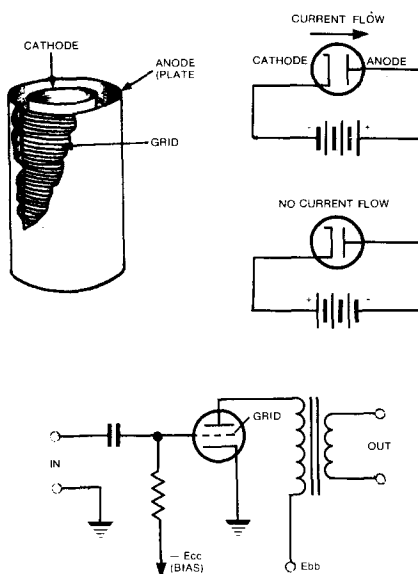Fig 1. Basic diode tube construction and operation.

Fig 2. Triode construction, circuit and characteristics.

electrons, but because of the grid's open structure, most of them continue on to the anode. Note the screen voltage takes precedence over the anode in controlling current. And we can swing the anode voltage further for more output, and get higher gain too.

The addition of the second grid with a fixed high potential results in a current flow essentially independent of anode voltage, but still subject to the action of the control grid. (Figure 3). Trouble occurs, however, when we try
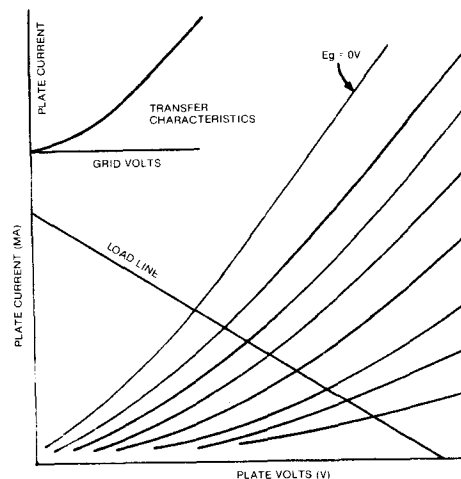
to produce an anode voltage swing lower than the screen voltage. Electrons are moving so fast that when they strike the anode they dislodge other electrons, which are attracted to the higher potential screen grid, thus reducing current through the load.
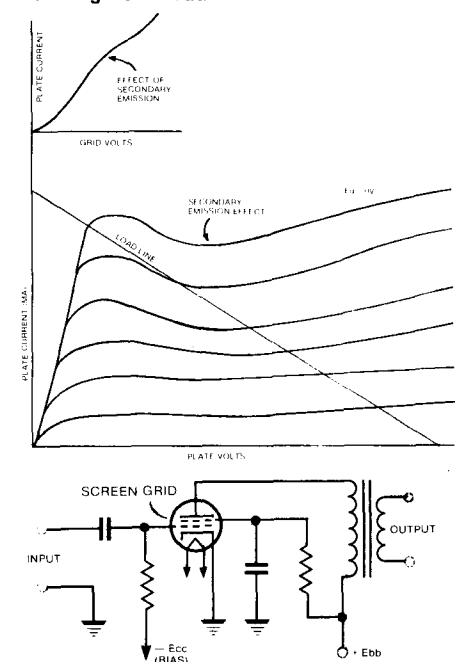

Fig 3. Series output arrangement.

VFETS FOR EVERYONE

This problem was overcome by adding a third grid between the screen and plate and tied to the cathode. Because it is at cathode potential the grid pushes the secondarily emitted electrons back to the anode, resulting in a family of curves as in Figure 4.

Distributed loading is also possible by dividing the load between screen and anode, and results in Figure 4a. This kind of flexibility makes it possible to design circuits of exceptional linearity.
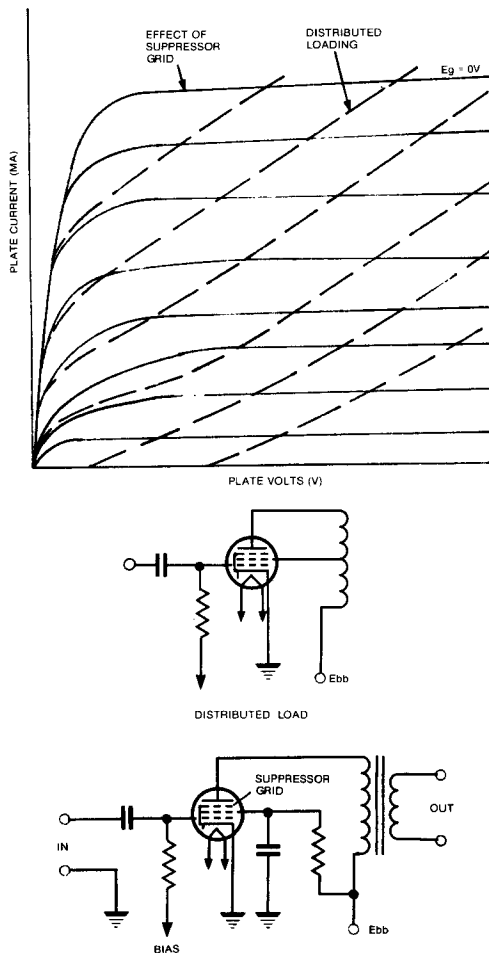

Fig 4. Single ended output with current source.

Problems

So far so good — except for a few problems. To begin with, the valve, like a light bulb, converts more electricity to heat than to useful work. It's very inefficient — for example the author (who is associated with the Canadian version of ETI) uses two 75 watt output class AB valve amplifiers to keep his studio at 25° C. without any additional heating in a Canadian mid-winter!

Also like a light bulb, a valve's performance deteriorates from the moment power is applied. Thus, direct coupled circuits can give real headaches in maintaining correct operating characteristics.

And then there's the output transformer. In order to match the thousands of ohms impedance to a low impedance load such as a loudspeaker, a transformer is virtually a necessity. With the inefficiencies already involved we can't afford the resultant impedance mismatches if we try to eliminate transformers. And we can't use gobs of feedback to reduce the resulting distortion. It's bad enough that, if we don't opt for a delicately balanced direct coupled circuit we have a low frequency roll-off and 90° phase shift at every R-C coupling point, but we have in any case the additional phase shift and internal resonances of the transformer. In practice, we are limited to between 20 and 26 dB of overall feedback. Obviously, a high level of open loop linearity must be designed into such an amplifier.

A great deal of engineering energy was spent designing output transformerless amplifiers, but few were successful, and those that were often created more problems than they solved.

Some legendary amplifiers *were* built using tubes. The Williamson, (I have one in daily use and it still sounds great), Quad, Leak Point One, MacIntosh Unity Coupled. The Quad, for example, delivered all of 15 watts — and was rock stable driving an electrostatic (Quad, of course) at live performance levels. Mac's drove a lot of disc cutters (at 60 watts) to produce discs which still sound spectacular.

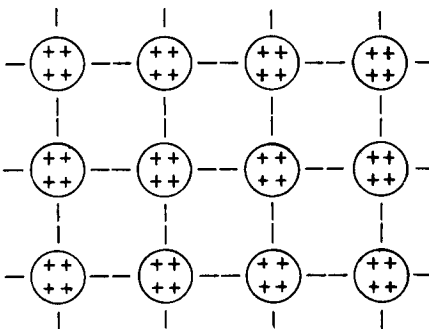But many were anxious to do something with the new-fangled transistors, and we did.

Fig 6. Drain to source resistance against temperature (Siliconix).

Transistors

The bi-polar transistor is composed of three materials, either a p-type semiconductor between two n-types, or an n-type between two p-types, (Figure 7a). A semiconductor such as silicon or germanium has a crystalline structure in the form of a diamond lattice with each atom having four adjacent neigh-

Fig 5b. P-type lattice structure.

Fig 5a. Basic lattice structure

Fig 5c. N-type lattice structure.

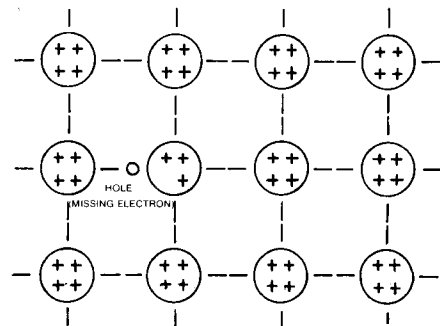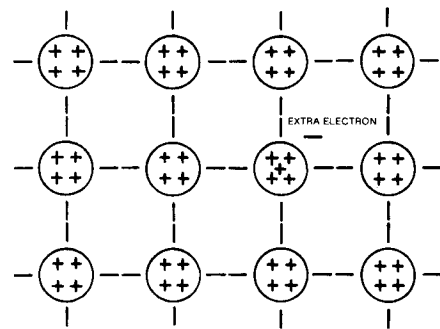bours, held together by co-valent bonds, each bond involving a shared pair of electrons. These electrons are not available for conducting current, so conduction is very semi. Indeed, resistance being around 100 million times that of copper.

However, if we introduce an impurity such as phosphorus or arsenic which has five valency electrons four of which form bonds while the fifth is only lightly held and is available for conduction. This is an n-type material (negative as it has an excess of electrons). If we add an impurity such as aluminium, only three valence electrons are available. Therefore, one of the valence bonds is not completed, resulting in a vacancy or hole in the lattice structure (Fig. 5). An electron from an adjacent electron pair bond may absorb enough energy to break its bond and fill the hold. This is a p-type material. This doesn't look like much of a big deal, but the result is quite dramatic.

Note that the atomic structure is in equilibrium — there is no net charge. However, if a free electron breaks its bond, it leaves behind a positive net charge; if it completes a bond by entering a hole, a negative net charge results. Current flow is produced by bringing about this carrier mobility. What was originally a very high resistance is now, under the right conditions, able to conduct substantial current, just as a small impurity (e.g. sulphuric acid) added to non-conductive pure water, makes electrolytic conduction possible.

When p and n-type materials are joined together, a p-n junction is formed (Fig. 6). Some of the free electrons from the n-type material diffuse across the junction and recombine with holes of the p-type material. The opposite process takes place with holes from the p-type material, producing a space charge or depletion region on either side of the junction, giving the p-type material a slight negative charge, and the n-type a slight positive charge. This process is finally limited by the resulting potential gradient.

If a battery is connected, as shown in Figure 6a, free electrons from the n-type material are attracted to the positive terminal, while holes from the p-type material are attracted to the negative terminal, widening the space charge region and increasing the potential gradient until it approaches that of the external battery. There is now little or no voltage difference across each region and little or no current flow. The junction is reverse biased.

If we reverse these polarities (Fig. 6b) electrons in the p-type material break their bond and enter the battery
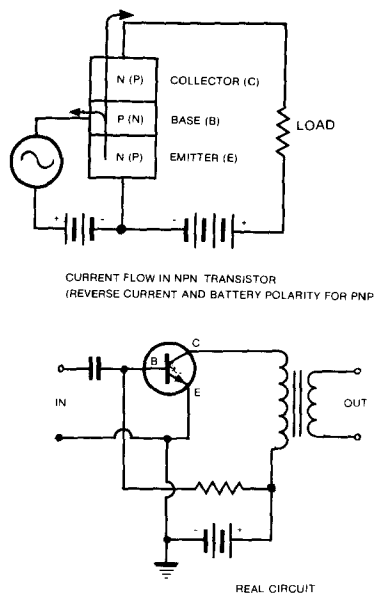

Fig 7a. Current flow in a semiconductor and circuit diagram.

creating new holes, while electrons from the battery negative terminal enter the n-type material and diffuse toward the junction. The space charge region narrows and the energy barrier becomes insignificant, so that excess electrons from the n-type material can penetrate the junction and move via the p-type holes to the positive battery terminal, for as long as voltage is applied. The junction is now forward biased.

Work!

In the device shown in Figure 7, the forward-biased emitter-junction injects electrons into the base region. The impurity or doping levels chosen are such that almost all the emitter current is composed of these electrons, and very few holes are injected into the emitter. The base region is very thin so that nearly all injected electrons diffuse to the edge of the depletion region of the reverse-biased base-collector junction where the field sweeps them across the collector bulk. Since for an equal current more power is developed across a high resistance than a low resistance, amplification occurs as a result of current being transferred from the low-resistance emitter-base junction to the high resistance collector junction.

The curves show that, as with the pentode tube, current is controlled mostly by the control electrode (base), but in this case the controlling parameter is current, not voltage. We have an inherently low-impedance device, and since it requires current into its input impedance, its signal source must be capable of delivering power. An ideal

transistor requires input current, unlike an ideal vacuum tube. This reduces efficiency but we don't have to heat up a cathode to shake a few electrons loose, so our overall efficiency is vastly greater.
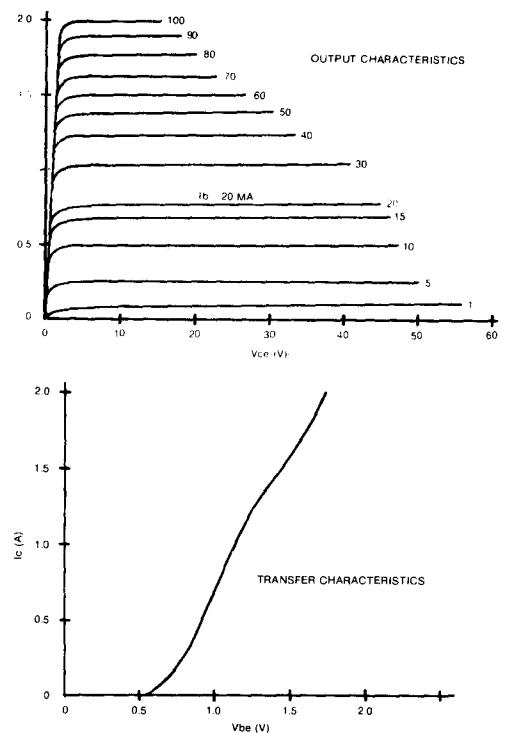

Fig 7b. Output and transfer characteristics of the 2N3054.

Disadvantage

The major disadvantage of this type of device lies in the nature of the depletion layers at the junctions, particularly the emitter-base. When current flows in a transistor, excess charge is stored in the base region. If the base-emitter junction is changed from a forward to reverse bias state, as in the negative swing of a class B or AB stage, or when a class A stage is overdriven, the junction cannot immediately switch to the reverse blocking state due to the presence of these excess charge carriers. They have the effect of allowing current to flow in reverse as if forward biased, until these charge carriers are removed.

In addition, there is capacitance effect associated with the barriers of a reverse-discharge time. The result is a switching transient during part of a cycle, sometimes erroneously referred to as crossover distortion (the latter occurs in any device in push-pull and is due to a discontinuity in the transfer function, usually caused by incorrect bias). This can be reduced by reducing the junction area but this reduces the dissipation capability. In fact, a transistor design favouring one characteristic usually does so at the expense of others.

Also, as temperature rises in the

VFETS FOR EVERYONE

device (due to current flow, for example) carrier mobility at the junctions increases, causing further increase in current. The current increase further raises temperature, which raises current — which further raises temperature — and so on. The resulting thermal runaway can quickly destroy the device. In milliseconds!

In large area transistors, current tends to become nonuniform in distribution. The temperature rise in the high current region leads to localized thermal runaway until equilibrium is reached by a sharp drop in collector voltage, (called secondary breakdown) frequently destroying the device. This is more true at high voltage and low current than the reverse, and frequently means that rated dissipation cannot be reached. This leads to overdesign, unnecessarily high voltage and dissipation ratings (and remember, a design which favours one characteristic often does so at the expense of others) plus elaborate protective circuits.

High levels of feedback are generally used to control distortion, and this in conjunction with the excess charge condition in the base, leads directly to transient overload, and resultant transient intermodulation. Output is delayed during this charge/discharge, which delays application of feedback. It simply isn't available. The input signal is not immediately reduced by feedback, and passes through at high initial level.

The millenium has not quite arrived after all!

The FET
Since a semi-conductor is precisely that, a battery connected across the ends of a p-type or an n-type bar will cause current to flow through the material, just as it does through a vacuum tube. We discussed earlier the characteristics of a pn junction. If, for example, a p-type material is joined to the surface of an n-type bar, located between the battery terminals, a pn junction is formed, and if this junction is reverse biased, a space charge or field is produced of opposite polarity which will inhibit current flow, just as the control grid inhibits current flow in a vacuum tube. Changing this reverse voltage causes a large current change, and amplification results.

A simple junction FET is shown in Figure 8. With a given drain-source voltage, maximum current flows at zero
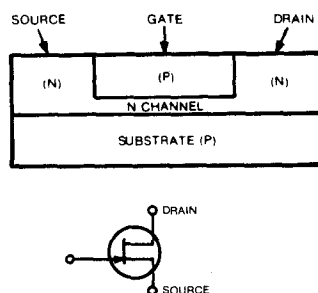

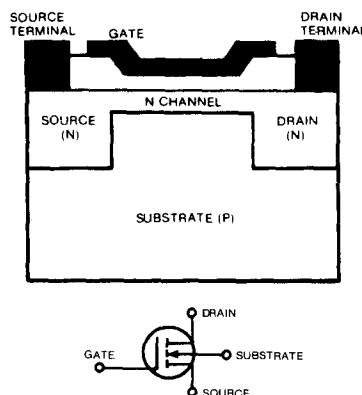Fig 8. N-channel JFET construction and symbol.

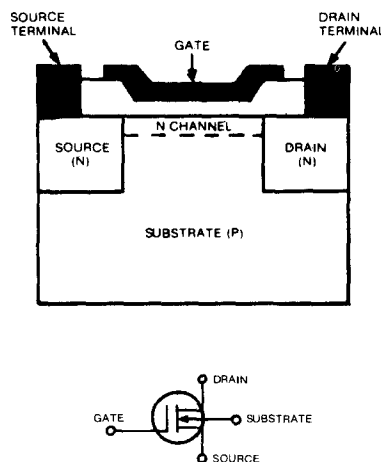Fig 9a. N-channel depletion horizontal MOSFET construction and symbol.

Fig 9B. N-channel enhancement horizontal MOSFET construction and symbol.

gate voltage, and at some reverse voltage, determined by device geometry and doping levels, no current will flow. Also, as in the vacuum tube, load characteristics are not reflected to the input circuit, because current is not controlled by carrier injection as in bipolars, but by voltage levels.

A variation is the Metal Oxide Semiconductor Field Effect Transistor. (MOSFET) (Fig. 9) a far more versatile device whose technology is virtually the cornerstone of modern computer technology, although it has had less use to

date in linear applications such as audio amplification.

MOSFETS come in two basic types. In both the gate consists of a metal electrode separated from the channel by a thin oxide layer. In the depletion type current flow is controlled by the electrostatic field of the gate when biased. When a depletion MOSFET is so biased the device may be driven on both sides of the zero volts point as with vacuum tubes. Unlike vacuum tubes, under these conditions, the gate draws no current, therefore does not require the driver to deliver power.

The enhancement type MOSFET shown in Figure 9b, is more widely used. The source and drain are separated by a substrate of opposite material, and under zero gate volts no current flows. However, when sufficient forward bias is applied to the gate the region under the gate changes to its opposite type (e.g. p-type becomes n-type) and provides a conductive channel between drain and source. Carrier level, and conduction is controlled by the magnitude of gate voltage.

Although MOSFETS are handy devices they are not capable of handling high power levels. The channel depth available for conduction is limited by the practical limits on gate voltage. The lower current density has been the primary limitation due to the horizontal current flow.

VFETS
Recent years have seen the introduction and commercial use of Vertical Channel J-FETS, notably by Sony and Yamaha (Fig. 10). The vertical channel permits a very high width-length ratio, permitting a decreased inherent channel resistance and high current density. Unfortunately it suffers the same disadvantages as the small signal J-FET, plus, in currently available devices, a very high input capacitance, ranging from 700 pF to around 3000 pF, limiting high frequency response. In addition, since they must be biased into the off condition, bias must be applied before supply voltage and removed after the supply if it is to be operated anywhere near its maximum ratings. This problem doesn't exist with vacuum tubes because of heater warm-up time, although some
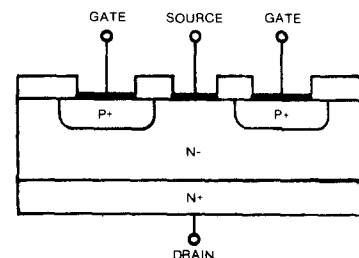

Fig 10. Vertical junction FET construction.

"instant-on" circuits impose heavy turn-on surges.

This necessitates a complex power supply, and Yamaha, for example, uses more devices in the supply than it does in its amplifier circuits. However, the construction does make possible the design of complementary types and both Nippon Electric and Sony have high power devices available. Unfortunately, neither company seems anxious to make detailed information available, so there is little to disclose here beyond the fact that they are said to have characteristics similar to those of triode tubes.

However, the Vertical MOSFETS by Siliconix are readily available, at reasonable prices, and the manufacturer most generous in providing data. The following information is extracted from their application note AN76-3, Design Aid DA 76-1, plus device data sheets.

The Device
Notice in Figure 11, that the substrate and body are opposite type materials separated by an epi layer (similar to high speed bi-polars). The purpose of this structure is to absorb the depletion region from the drain-body junction thus increasing the drain-source breakdown voltage. An alternative would have involved an unacceptable trade-off between increasing the substrate-body depth to increase breakdown voltage (but increasing current path resistance)
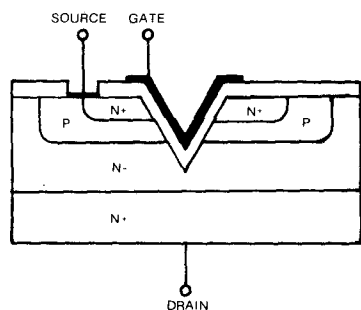

Fig 11. Vertical MOSFET construction (Siliconix).

and lengthening the channel. In addition, feedback capacitance is reduced by having the gate overlap n-epi material instead of n+.

In manufacture, the substrate-drain and epi layer are grown, then the p-body and n+ source diffused into the epi layer, in a simiar manner as the base and emitter of a diffusion type transistor. A V groove is etched through the device and into the epi layer, an oxide layer grown, then etched away to provide for the source contact and an aluminium gate deposited. This type of device allows current flow in one direction only; this is not always so with a similar type of horizontal FET, where source and drain may be identical in

structure and of the same material. Therefore, no reverse current flows (we hope) when used in switching applications, as was also the case with vacuum tubes.

In-circuit operation is refreshingly simple: Supply voltage is applied between source and drain, with the drain positive with respect to the source, under which conditions no current flows, and the device is off. This is an enhancement type device, and is turned on by taking the gate positive with respect to the source and body. The electric field induces an n channel on both surfaces of the body facing the gate, and allows electrons to flow from the negative source through the induced channel and epi and through the substrate-drain. The magnitude of current flow is controlled almost entirely by the gate voltage, as seen in the family of curves (Fig. 12) with no change resulting from supply voltage changes above 10 V.

Advantages
The vertical structure results in several advantages over horizontal MOSFETS.
1) Since diffusion depths are controllable to close tolerances, channel length, which is determined by diffusion depth, is precisely controlled. Thus, width/length ratio of the channel, which determines current density, can be made quite large. For example, the VMP1 channel length of about 1.5μ, as against a minimum of 5μ in horizontal MOSFETS, due to the lower degree of control of the shadow masking and etching techniques used in such devices.
2) In effect, two parallel devices are formed, with a channel on either side of the V groove, thus doubling current density.
3) Drain metal runs are not required when the substrate forms the drain contact, resulting in reduced chip area, and thus reduced saturation resistance.
4) High current density results in low chip capacitance. Also, unlike horizontal MOSFETS, there is no need to provide
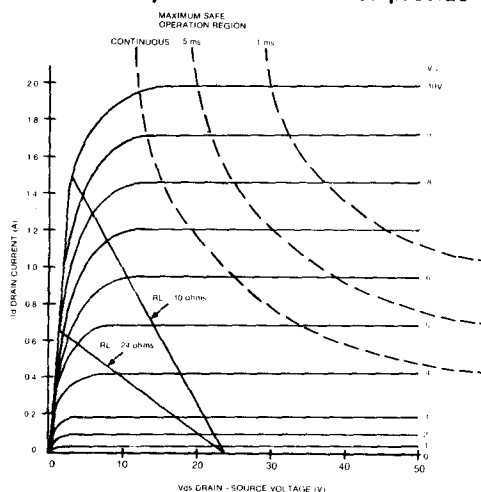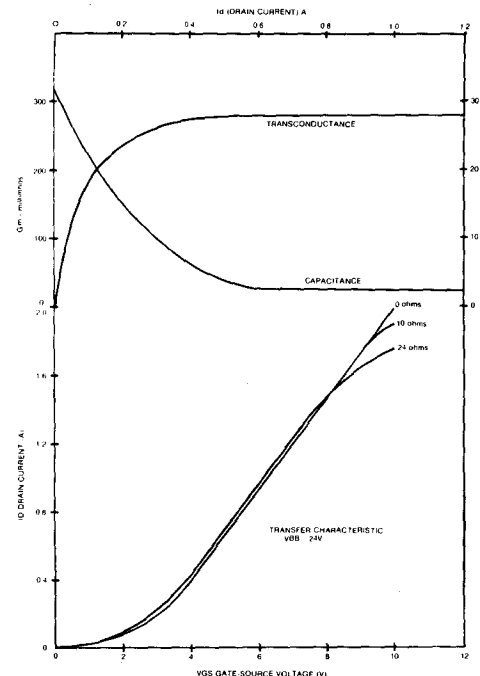

Fig 12a. Output characteristics VMP1.

Fig 12b. Other VMP1 characteristics.

extra drain gate overlap to allow for shadow mask inaccuracies, so feedback capacitance is minimized.

In comparison with bi-polars, especially power devices, the advantages are even more impressive.
1) Input impedance is very high, comparable to vacuum tubes, since it is a voltage controlled device, with no base circuit drawing current from the driver stage. A 7 V swing at the gate, at virtually 0A, represents almost 0W of power, but can produce a swing of 1.8 A in output current. This represents considerable power gain and will interface directly with high impedance voltage drivers.
2) There is no minority carrier storage time, no injection, extraction, recombination of carriers, resulting in very fast switching and no switching transient in class B and AB amplifiers. Switching time for a VMP1 is 4 ns for 1 A, easily 10-200 times faster than bi-polars, and rivalling many vacuum tubes.
3) No secondary breakdown, and no thermal runaway. VMOS devices exhibit a negative temperature coefficient with respect to current, since there is no carrier recombination activity to be speeded up with temperature. Thus, as current increases so does temperature, but the temperature rise reduces current flow. It is still possible to destroy the device by exceeding its maximum ratings, but a brief near-overload does not result in an uncontrollable runaway condition. Usually, simple fusing and/or thermistor protection is sufficient for maximum safety, and even this may be unnecessary with conservative design. Absence of secondary breakdown means that full dissipation can be realized even at higher supply voltages. In this respect they resemble vacuum tubes. ∎

Sooner or later you'll need a Gregory's.

VFETS FOR EVERYONE

In the first part of this article we examined the structure and features of a new type of semi-conductor, the vertical channel power metal oxide semi-conductor field effect transistor, Vertical MOSFET, or V-MOS recently introduced by Siliconix. The second part of the article covers the actual use of V-MOS.

V-MOS POWER FETS like signal MOSFETS, may be used to perform many different functions. However, no matter what the circuit, certain conditions, common to all applications, must be provided. These are supply power, loading, drive signal, and establishment of appropriate operating points.

The electrical characteristics of the VMP1, VMP11, and VMP12, are shown in Fig. 1, and Fig. 2 shows them in graphic form. Since these are uni-directional devices, the source and drain are not interchangeable, and as they are n-channel devices conduction can occur only if the drain is positive with respect to the source, and high enough to ensure operation in the linear region — as with a vacuum tube, bi-polar transistor, or signal FET.

Like the vacuum tube, the absence of secondary breakdown allows full dissipation at any voltage supply up to maximum voltage and current ratings.

Thus, where two different designs require the same dissipation but different voltage/load current, no derating is required. This is shown in the "safe operating area" curves. The only bi-polar transistor possessing this characteristic is the single-diffused type, which is also the least suitable for any application requiring wide bandwidth and/or high speed.

This characteristic also simplifies the establishment of suitable load-lines allowing greater safety margin in driving reactive loads where the load-line may be elliptical to the point of leaving the safe-operating area. Designers accustomed to using high voltage high dissipation devices to assure adequate safety margins at relatively low power levels need not therefore be too disconcerted at the 25 watt rating of these devices.

A 10 watt class A amplifier suitable for driving a tweeter in a bi-amped speaker system, for example, need not suffer excessive dissipation except perhaps with an electrostatic unit where such a power level would be inadequate anyway, unless it were operating at a very high cross-over frequency.

Output

These devices may use any of the types of output circuits in general use with tubes and bi-polars, including transformer coupled (Fig. 12) where the benefits of the absence of charge carrier storage become apparent in the absence of severe ringing at the cross-over point, conventional series output such as in Fig. 3, which is a straight-forward transformation from a bi-polar

Fig. 2 Typical VMP1 performance curves (Siliconix).

Fig. 1 Electrical characteristics of the VMP devices (Siliconix).

		Characteristics		VMP 11			VMP 1			VMP 12			Unit	Test Conditions
				Min	Typ	Max	Min	Typ	Max	Min	Typ	Max		
1		BVDSS	Drain-Source Breakdown	35			60			90			V	$V_{GS}=0$; $I_D=100\,\mu A$
2	S T A T I C	$V_{GS(th)}$	Gate Threshold Voltage	0.8		2.0	0.8		2.0	0.8		2.0		$V_{GS}=V_{DS}$; $I_D=1\,mA$
3		I_{GSS}	Gate-Body Leakage			0.5			0.5			0.5	μA	$V_{GS}=15\,V$; $V_{DS}=0$
4		$I_{D(off)}$	Drain Cutoff Current			0.5			0.5			0.5		$V_{GS}=0$; $V_{DS}=24\,V$
5		$I_{D(on)}$	Drain ON Current*	1	2.0		1	2.0		1	2.0		A	$V_{DS}=24\,V$; $V_{GS}=10\,V$
6		$I_{D(on)}$	Drain ON Current*	0.5			0.5			0.3				$V_{DS}=24\,V$; $V_{GS}=5\,V$
7	S W I T C H	$r_{DS(on)}$	Drain-Source ON Resistance*		2.0	2.5		3.0	3.5		3.7	4.5	Ω	$V_{GS}=5\,V$; $I_D=0.1\,A$
8					2.4	3.0		3.3	4.0		4.6	5.5		$V_{GS}=5\,V$; $I_D=0.3\,A$
9					1.2	1.5		1.9	2.5		2.6	3.2		$V_{GS}=10\,V$; $I_D=0.5\,A$
10					1.4	1.8		2.2	3.0		3.4	4.0		$V_{GS}=10\,V$; $I_D=1\,A$
11	D Y N A M I C	g_m	Forward Transconductance*	200	270		200	270		170			m℧	$V_{DS}=24\,V$; $I_D=0.5\,A$
12		C_{iss}	Input Capacitance		48			48			48		pF	$V_{GS}=0$; $V_{DS}=24\,V$ $f=1\,MHz$
13		C_{rss}	Reverse Transfer Capacitance		7			7			7			
14		C_{oss}	Common Source Output Capacitance		33			33			33			
15		t_{ON}	Turn ON Time**		4	10		4	10		4	10	ns	See Switching Time Test Circuit
16		t_{OFF}	Turn OFF Time**		4	10		4	10		4	10		VMC

*Pulse Test   **Sample Test
Pulse Test Pulse Width = 80 μsec, Duty Cycle = 1%
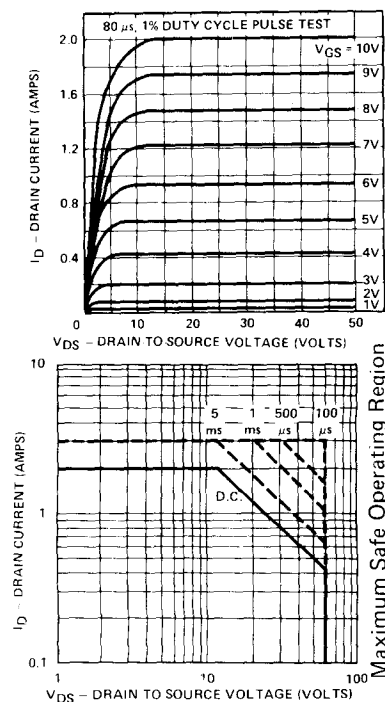
VFETS FOR EVERYONE

circuit (1), and single-ended output with current source, also transposed from an excellent bi-polar circuit (2) (Fig. 4).
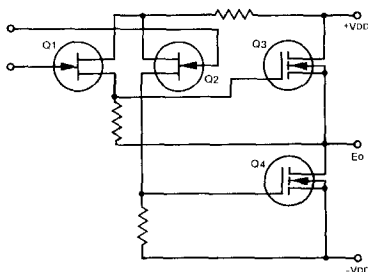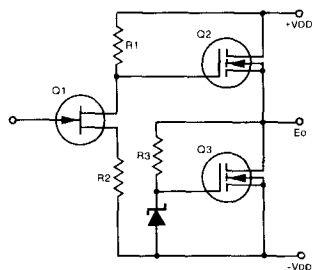

Fig. 3 Series output arrangement

Fig. 4 Single ended output with current source.

Bias and Drive

These series of devices are n-channel, enhancement type MOSFETS, and may be biased and driven using methods appropriate to signal types and bi-polars. The drain is made positive with respect to the source and the gate enables conduction by being forward biased with respect to the source, that is to say it is biased in a positive direction. Unlike bi-polars, however, they are voltage, rather than current controlled, and circuit values are selected to provide the required voltage. Any current drawn is by the bias network itself.

Three bias methods are shown in Fig. 5. Figure 5a shows bias supplied from a fixed bias supply. It is the simplest possible method, allows extremely high input impedances since Rg may be almost any very high value desired, and its stability is limited only by the stability of the bias supply.

The design shown in Fig. 5b has the advantage of requiring no extra supply voltage since it is taken from Vdd. Disadvantages are low impedance and

stability. Input impedance consists of the parallel combination of R1 and R2 (disregarding input capacitance of the MOSFET and the very low input leakage.) There are practical limits as to how high this combination can become; if for example, we have a 60 volt supply and require 6 volts bias, we might have some difficulty obtaining higher values than 9 megohms and one megohm for R1 and R2.

Higher values become more difficult to obtain, stability becomes less reliable, internal inductance and distributed capacitance become problems, and overcoming these difficulties usually costs money. In addition, if Vdd is subject to variation, then bias varies. In a class AB amplifier this could be quite serious, since Vdd varies considerably with output level; at high levels, Vdd can be expected to drop, causing a reduction in bias.
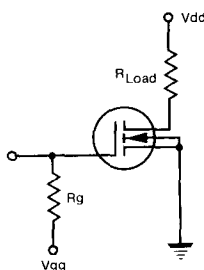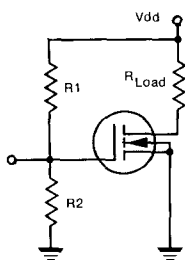

Fig. 5a. Hi-Z separate bias supply.

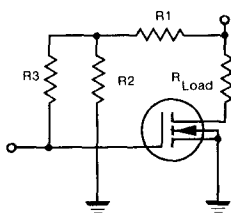*Fig. 5b. Moderate impedance supply.*

Fig. 5c. Hi-Z common supply.

While this may reduce the danger of over-driving the device, it will be forced to operate in its non-linear region which may result in unacceptable performance characteristics unless taken into consideration in the overall circuit design (e.g.

choice of feedback values). It does provide some degree of overload protection, and with correct choice of values can provide for class AB operation at low levels, shifting to class B at high levels. With these considerations in mind, and/or where moderate impedances are required, it offers a low cost, simple, and reasonably reliable method of establishing the operating point.

The method used in Fig. 5c is similar except that with the addition of R3 higher input impedances are possible. Its configuration is similar to a noiseless biasing system frequently used in low-level bi-polar amplifiers and integrated circuits (e.g. National LM381A) but its function is somewhat different. Resistors R1 and R2 form a voltage divider as in Fig. 5b, but their junction now forms a fixed bias source as in Fig. 5a. Resistor R3 can be quite high since no current flows. Meanwhile, since the parallel combination of R1 and R2 are effectively in series with R3 they can be reduced to more manageable values. Alternatively R2 can be replaced by Zener diode for stability comparable to Fig. 5a.

Input Protection

Unlike most signal MOSFETS, the gate of each of these devices, with the exception of the VMP4, is protected with an internal 15 volt, 10 mA zener diode. Most signal MOSFETS, as well as the VMP 4, are unprotected, or where extremely high impedances are not required, are protected by back to back zeners. I have no information as to why this different technique is used.

This different technique is used, but it is obvious that a negative signal swing on the gate will result in forward current through the zener. If the device is to be driven beyond cutoff, the driver must be capable of delivering current during its negative swing. Alternatively, a constant current source can be used, a series limiting resistor or a driver biased to the same class of operation as the V-MOS FET.

A constant current source (we'll examine an example of its use a little later) will limit current drive to the value of the constant current diode used; a series resistance will drop the drive voltage as the diode draws current. In both cases, diode current must be limited to 10 ma maximum. Higher currents will damage the protective zener diode.

In amplifier applications, a class A driver is commonly used. However, if a class B output is used, conduction only occurs during positive half-cycles. Therefore drive signal is not required during negative half-cycles. If a source or

emitter follower driver stage is biased so as to pass no negative drive, the problem does not occur. However, great care must be exercised in the design of such a stage to ensure that drive does not disappear before the output device is cut off.

This is not too difficult with a class B or near class B stage; If the output device is operated at zero bias, then a small amount of bias on the driver will ensure conduction during slightly more than 180 degrees. Class AB operation is a little more tricky. If conduction is to occur for 270 degrees, for example, the driver should conduct for slightly more than this period.

Two types of drive circuits familiar to designers of bi-polar circuits are the Darlington and super beta, commonly used together to provide a quasi-complementary circuit. Both circuits are current amplifiers designed to provide a compound device with very high hfe and provide base current to the output device. However, similar circuits can be used with these devices to provide phase inversion in a series output stage.

Thermal considerations

As described earlier (Part 1) these devices exhibit a negative temperature coefficient with respect to current, so that as temperature rises, current is reduced, thus providing a self-inhibiting action which provides some protection against overload. However, this is not an unconditional effect Fig. 6 shows the relationship between RDS(on) and temperature (3), based on a worst case temperature coefficient of 0.7 per cent per degree C.

Suppose that the device when 'on' passes a current of 1 amp which causes it to heat up. The 'on' resistance increases (which is why current drops), increasing the voltage drop across the device and the device dissipation. Now, if adequate heat sinking is used there is no real problem but if it isn't, the 'on' resistance and junction temperature will rise to the point where extra charge carriers are generated, thus stabilizing
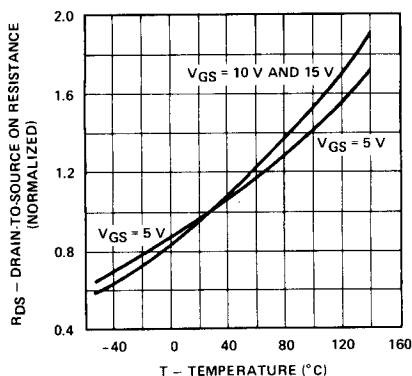
Fig. 6 Drain to source resistance against temperature (Siliconix).

RDS(on). That's great, except for the fact that this doesn't occur until the maximum safe junction temperature of 150 degrees has been exceeded.

You'll remember that we said earlier that the device was free of thermal runaway problems because of its negative temperature coefficient, but it isn't free of thermal destruction problems, and in any case, excessive temperatures will reduce output conductance. Heat-sinking requirements are, therefore, similar to those of bi-polars. The calculations of thermal operating conditions are beyond the scope of this article, but interested readers are referred to the Siliconix literature listed in the references, (4).

Extending the ratings

The current handling and therefore total dissipation capability may be increased by simply connecting several devices in parallel (Fig. 7). No ballast resistors are needed to ensure proper current sharing since if one device draws more current than another it simply gets a little warmer which causes it to draw less (assuming adequate heat sinking, of
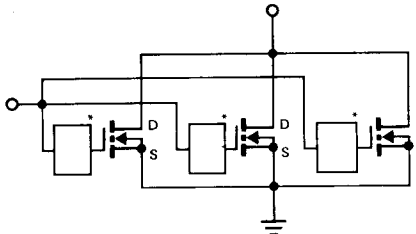
*TO PREVENT SPURIOUS OSCILLATIONS, A 500 Ω - 1K Ω RESISTOR OR FERRITE BEAD (FOR HIGHER SPEED) SHOULD BE CONNECTED IN SERIES WITH EACH GATE.

Fig. 7 Basic circuit for parallel operation (Siliconix).

course). The only major precaution needed is to keep lead inductance in the gate and source connections to a minimum to prevent parasitic oscillations, unless the devices are driven from a low impedance source.

It may be advisable to insert what the British call "stoppers" — small resistors (100 to 1000 ohms) in series with each gate, wired directly to the socket, or ferrite beads mounted on the leads close to the socket terminals. An additional plus when paralleling several devices is that the gm is multiplied by the number of devices used. Mutual conductance gm is specified as the ratio of a large change in current to a small change in control voltage. If, for example, a change of 0.4 volts on the gate produces a change of 0.1 amp through one device, connecting two devices in parallel will give us an output swing of 0.2 amps, but it will still require only the original 0.4
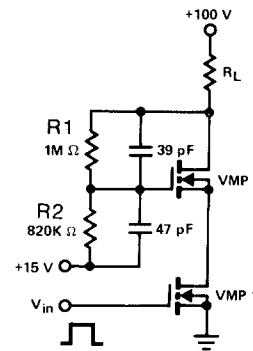
Fig. 8 Diagram for series operation (Siliconix).

volts gate swing. Since voltage gain A=gm x RL, if gm is increased, A is increased.

In real use, of course, the internal resistance of two devices in parallel is less than of one, the optimum load is less, so in amplifier applications, the net amplification A is the same. But notice that the drive requirements have not changed. With bi-polars current would have to be supplied to each base, thus increasing the output requirements of the drivers. Indeed, with many high-power amplifiers using multiple output devices the drivers are also power devices.

We can also extend the voltage ratings by series operation of two or more devices Fig. 8 shows the technique. Resistors R1 and R2 bias Q2 'on', while C1 and C2 ensure fast switching. The input control signal is inserted between gate and source of Q1. Ordinarily the bottom of the divider chain is at ground potential for signal frequencies, so that circuit is really a cascade.

Maximum current and gm are the same for one device.

Some practical applications

An efficient light dimmer circuit as proposed by Siliconix is shown in Fig. 9. The 4011 acts as a pulse width modulated oscillator whose duty cycle is determined by the ratio of R1 to R2, with R2 adjusted to control the brightness of the W-90 bulb. Of special interest here is the fact that with its fast switching time, the VMP1 is especially suited to pulse width modulation at power levels and suggests it as being suitable for use in switching, or class D linear amplifiers.

A DC to DC converter is outlined in Fig. 10. The VMP1s form an oscillator with positive feedback provided by the additional coil in the gate circuits. In operation the upper V-MOSFET is biased 'on', and the lower V-MOSFET
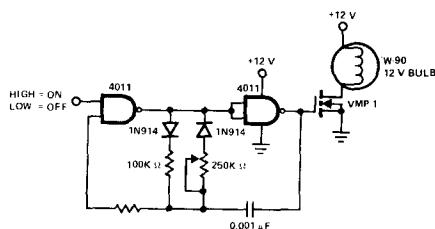
VFETS FOR EVERYONE

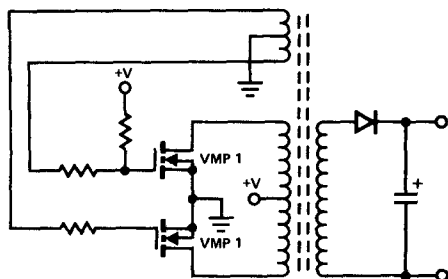Fig. 9 Circuit of a high efficiency light dimmer (Siliconix).

Fig. 10 A d.c. to d.c. converter (Siliconix).

is 'off'. When power is applied the upper device conducts causing current to flow from Vdd through the upper half of the transformer primary and the upper V-MOSFET to ground. The induced current flow through the feedback coil develops a voltage such as to shift the bias in the upper device 'off' (if the winding is connected with the correct polarity) and the lower device 'on'. This causes current flow from Vdd through the lower half of the transformer primary and the lower V-MOSFET to ground.

The secondary circuit consists of a single rectifier and filter. The resistor in the upper gate prevents shorting out gate bias, and the one in the lower gate keeps both sides balanced. In addition, each resistor limits current through the protective diodes. These are expen-

sive devices for such an application, but the high reliability, the reduced rf radiation (due to reduced switching transients) and the circuit simplicity easily make up for the cost. The very high circuit impedance enables running frequency to be set by the self-resonance of the transformer.

Single-ended push-pull transformer coupled audio amplifiers are shown in Figs. 11 and 12. Both utilize the biasing system described in Fig. 5b. A loadline drawn on the output characteristic will show the optimum load to be 24 ohms. In Fig. 11 gate drive is supplied by a single junction FET, and voltage feedback is taken from the output transformer secondary and series fed to the source of the input device. Distortion is under 2% at full output (try to get that with a single ended tube or bi-polar) and could probably be reduced even further by adopting a source follower output stage.

A push-pull version of Fig. 11 is shown in Fig. 12 using a differential input to provide phase splitting, drive, and a feedback point. Although the transformer winding ratio implies the use of a low impedance loudspeaker, a step-up ratio could be used for direct coupling to an electrostatic speaker, a balanced transmission line (both with some modification of the feedback circuit) an unbalanced transmission line, or a 70 volt speaker distribution line.

Notice in both circuits, and in the biasing circuits of Fig. 5, that no source resistors have been used, either for local feedback or for bias setting. In tube and bi-polar circuits it's a useful technique, and with bi-polars can be used to stabilize bias and control thermal runaway by using the increased current flow to increase the voltage drop, thus reducing base-emitter voltage. However, if used with these devices, it will actually impair the self-limiting action of its negative temperature co-efficient. If temperature rises due to high current, current flow is reduced. This would reduce the voltage drop across a source

resistor, lowering the source voltage and increasing the gate-to-source voltage, causing an increase in current flow. The circuit would work great while it lasted — which wouldn't be for long.

Record amp
Figure 13 shows a magnetic recording amplifier derived from a tube circuit. Its biggest advantage lies in its ability to provide equalization for head losses by incorporating the head within the feedback loop. Additional equalization is then required only for gap losses and tape self-demagnetization. Q1 acts as a driver for Q2, the output stage, which, with series resistor R9, provides a high impedance current source for the record head, as well as providing a mixing pad between audio and bias currents.

The record head's return path to ground is through R11. The inductance of the record head results in an impedance characteristic which rises with frequency. At frequencies at which the impedance of the head is low in comparison with R9 and R10 in series, load current is essentially constant. As frequency rises, however, head impedance becomes appreciable. With appropriate selection of R9 and R11, depending on head characteristics, the voltage across R11 decreases as the head impedance becomes significant. If feedback is taken across R11 it will decrease with rising frequency, causing an increase in gain, at a rate of 6dB/octave.

Feedback is applied across R3 via R10 and C8 (which supplies bass boost below 80 Hz) C5 and C6 provide additional high frequency boost for a total ultimate slope of 12 dB/octave. This circuit is so effective that no additional boost is needed at 15 ips, and only a small amount at 7.5 ips with high coercivity tape.

The biasing method used is that of Fig. 5c. The large amount of local current feedback provided by R2 and R3 results in a high output impedance
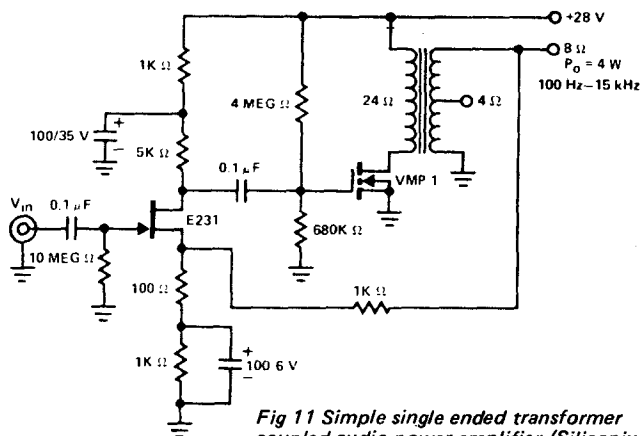
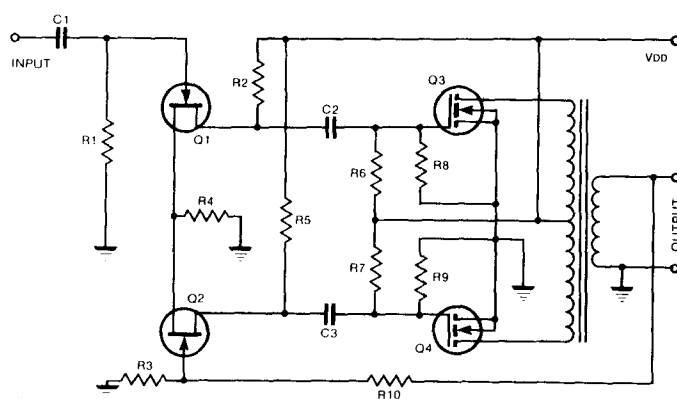Fig 11 Simple single ended transformer coupled audio power amplifier (Siliconix).
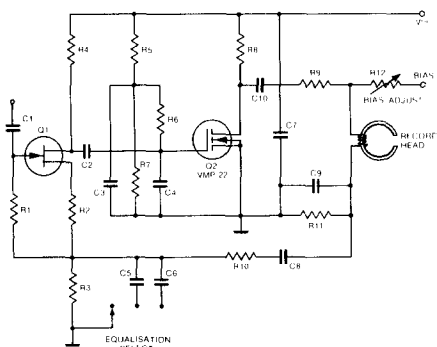
Fig. 12 Transformer coupled output.

Fig. 13 Tape recording amplifier.

for Q1, so the biasing network is selected to provide high impedance with reasonable values. Capacitors C3, C4, C7 and C9 bypass bias signal to avoid overloading Q2, and to prevent attenuation of bias current.

Power amp

Figure 14 shows a high quality power amplifier designed by Siliconix Inc. (5) and described in their application notes. Output current capability is increased by using three VMP12's in parallel, providing for 6 amp current 75 watt dissipation and load optimized at 8 ohms. Q11-13 operate as a source follower, while Q8-10 form a quasi-source follower. This is accomplished by applying local feedback from drain to gate via R14, R15, and driving the gate by a modified current source. This consists of a cascade circuit with a constant current diode as the load.

For the benefit of those not familiar with these devices, a constant current diode is really a FET connected internally as shown in Fig. 15. Since current in a FET is controlled essentially by the gate-to-source voltage, changes in load

or in applied drain-to-source voltage have negligible effect since gate-to-source voltage is held constant. This is a current analogue to the zener diode and is described in detail in Siliconix literature (6).

The design is push-pull from input to output, thanks to differential circuitry throughout, prior to the drivers. Open loop distortion is low, bandwidth wide, allowing satisfactory performance with only 22 dB of feedback. Lead compensation only is used (via C4), along with the liberal use of local feedback (R4, R5, R11, R12,). The result is very low transient IM and a slew rate of over 100 V/microsecond. THD is quite respectable even though the numbers might not impress the average audiophile accustomed to amplifiers with great specs and poor sound.

Incidentally, D8 and D9 illustrate an excellent method of providing output current limiting. In this case, 9.1 volt zener diodes limit drain current to slightly less than 2 amps. At first one might be tempted to depend on the built-in protection diodes to accomplish this, but it should be remembered that these devices are for protection against static discharge. Their zener voltage of 15 volts at 10 mA cannot possibly be used since the absolute maximum permissable drain current occurs at a gate-to-source voltage of 10 volts.

Commercial amps

A simplified version of Yamaha's B1 amplifier is shown in Fig. 16, (8). In this circuit a cascade drive system is used, but in a differential form with the constant current source in the

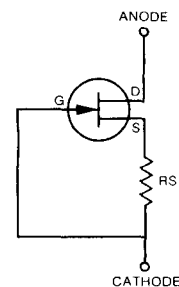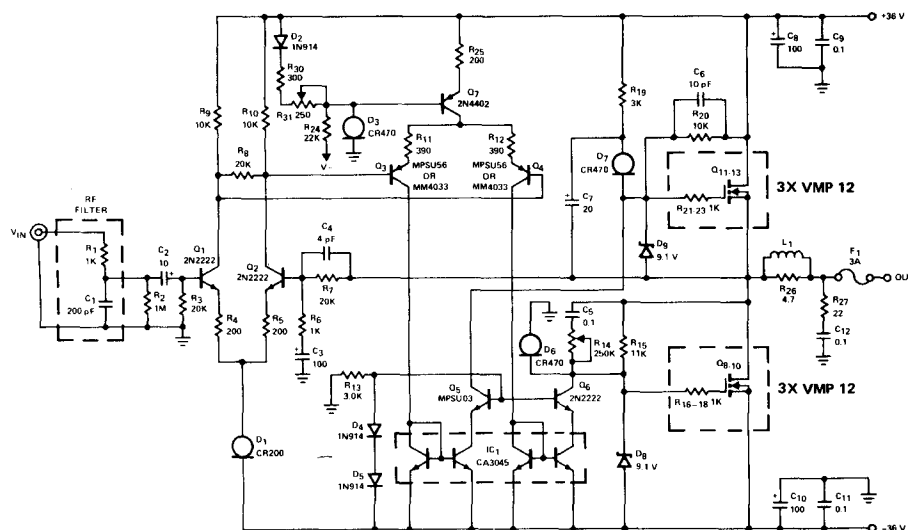*Fig. 14 A high quality 40W amplifier (Siliconix).*

Fig. 15 A FET as a constant current diode.

common source circuit. This is an example of all FET design of excellent performance and received rave reviews in several publications including ETI. It's also inexpensive!

The VHF linear amplifier in Fig. 17 will deliver 5 watts peak envelope with second and third order intermodulation at -30 dB from 144-146 MHz. It will also prove useful as a receiver pre-amp with a noise figure of 2.4 dB. V-MOSFETS show considerable promise in rf applications because of their linear transfer characteristic, the high gain capabilities even with Ft somewhat above 600 MHz, low noise and (in receiver front ends) very wide dynamic range. Although this article has dwelt on the VMP 1 family, there is also the VMP 4, designed specifically for rf applications and which is now available.

Finally, how about something elegant for its simplicity, such as the tapered current voltage limiting battery charger shown in Fig. 18. This is especially useful with Ni-Cad batteries which are intended for stand-by use and are permanently on charge, such as electronic clocks. Overnight shut-downs of a few hours are occasionally but irregularly experienced. You know what this can do to clocks. Especially alarm clocks which are supposed to make noises, turn on radios, start the coffee at a pre-set time in the morning so you can go to work. Battery operation is not too satisfactory if the readout is on continuously, and Ni-Cads should not be on permanent floating charge.

With this little device current is supplied to the battery via the VMP-1. Gate voltage is set at a value equal to the desired end-of-charge voltage. As the battery charges, its voltage increases, reducing gate-to-source voltage, thus reducing charging current. When the battery reaches full charge its voltage, and that of the source, equals gate voltage, and charge is terminated. If a load is placed across the battery it will draw current, and as the battery voltage drops slightly below gate voltage, charging at a trickle rate occurs — automatically.

VFETS FOR EVERYONE

Experimentation

The various applications shown are intended as suggestions for further experimentation. They are mainly designed to illustrate various characteristics of the device under consideration, and are not necessarily representative of commercial practice or of finished designs. In some cases this may be just as well! But we would be delighted to hear of any readers' experience with any of these or other circuits.

The author's feeling is that V-MOS constitutes a genuine breakthrough in semi-conductor technology, as important as the silicon transistor and the FET itself. We'll be seeing more of these devices, with higher ratings (a 10 amp 200 volt unit is already under development) and specialized characteristics. They are said already to be in use commercially as magnetic core drivers.

Digital enthusiasts may be somewhat impatient with the strong emphasis on audio applications in this piece but other literature has placed great emphasis on digital applications, with little attention paid to linear techniques beyond the 40 watt amplifier described here. The serious reader in all areas is referred to the references at the end.

Further literature may be obtained from Siliconix Inc., 2201 Laurelwood Rd., Santa Clara, CA 95054. They have been most helpful in providing information for the preparation of this article.

Have fun.
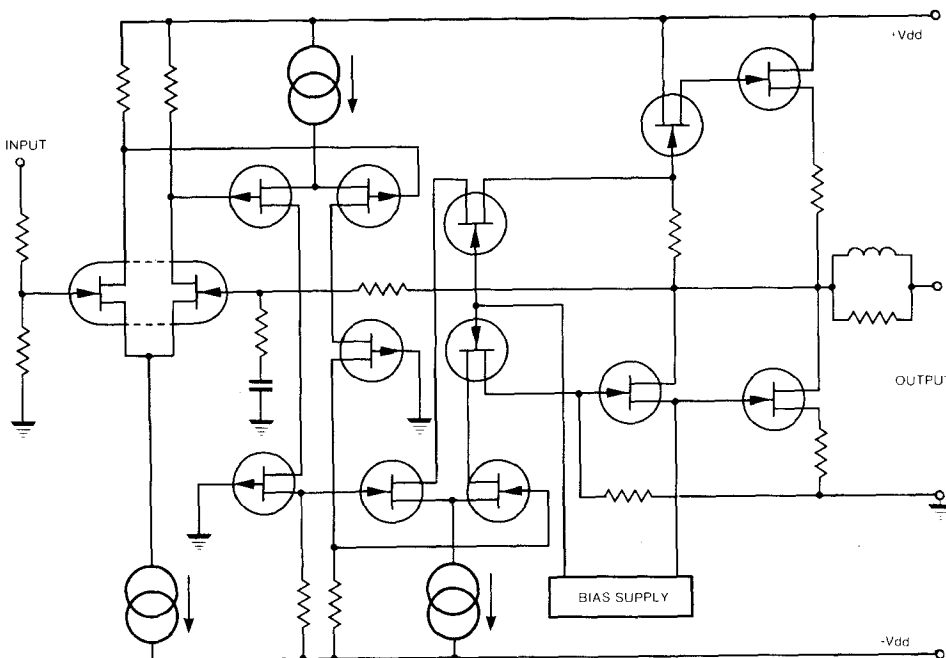
Fig. 16 Simplified Yamaha VFET amplifier diagram.
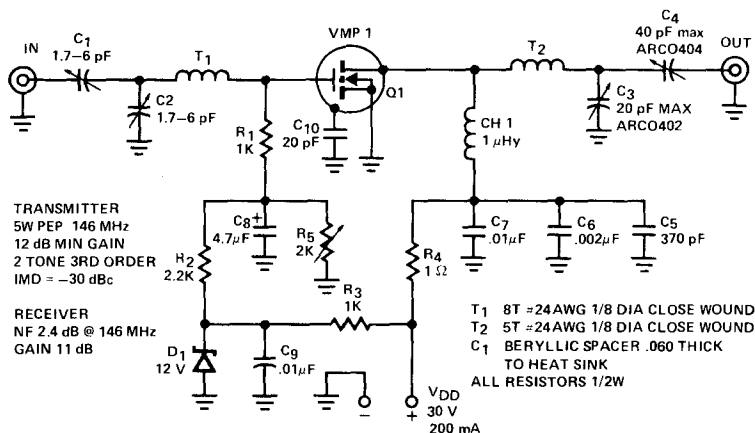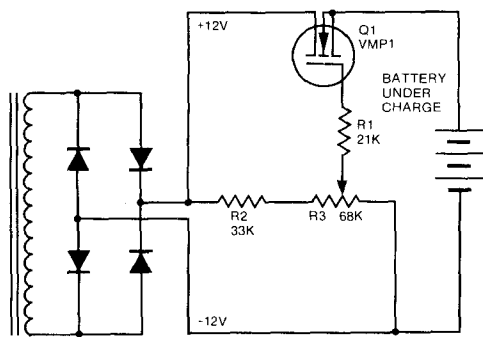

Fig. 17 144-146 MHz linear amplifier (Siliconix).

Fig. 18 Tapered current voltage limited battery charger.

REFERENCES

1. W. Marshall Leach: "Construct a Wide Bandwidth Preamplifier" Audio, Feb. 1977, p.39.
2. Nelson Pas: "Build a Class A Amplifier" Audio, Feb. 1977, p. 29.
3,4 Lee Shaeffer: "VMOS-A Breakthrough in Power MOSFET Technology" Application Note AN76-3, May 1976, Siliconix Inc.
5. Lee Shaeffer: "The MOSPOWER-FET Audio Amplifier" Design Aid DA76-1, May 1976, Siliconix Inc.
6. "The FET Constant Current Source" Design Idea D171-1, January 1976. Siliconix Inc. Also Siliconix Field Effect Transistors Data Book.
7. Ref. 5.
8. Bascom H. King: "Power FETs" Audio, Feb. 1975, p. 42.
10. Ed Oxner. "A New Technology: Application of MOSPOWER FETs For High Frequency Communications" Technical Article TA76-2, Nov. 1976, Siliconix Inc.

Power MOSFETS — the technology, the techniques

Brian Dance

IN 1976 Siliconix startled the semiconductor world with a new type of power MOSFET device. Recently other manufacturers have produced many other types of MOSFET products which are challenging power transistors and Darlingtons.

The name MOSFET stands for Metal Oxide Silicon Field Effect Transistor. Field effect transistors (FETs) are essentially voltage controlled devices, unlike conventional transistors in which the small base current controls a larger collector output current. FETs have very high input impedances so that very little input current is required to control their output current.

The input impedance of MOSFETs is especially high because they have an insulating film of silicon dioxide between the input gate electrode and the channel through which the output current flows. The gate electrode is therefore essentially completely insulated and virtually no input current can flow.

Various types of small MOSFET devices have been available for many years. Internally they contain a very small silicon chip on the surface of which the MOSFET device has been fabricated. Any current passes through these devices in a horizontal direction through the very thin surface layers and therefore the maximum current is quite low; maximum power dissipation in such devices is not normally over 1W.

VMOS devices

In the so-called VMOS devices, developed by Siliconix about four years ago, the current flows vertically through the semiconductor material — hence the name VMOS. This name is also associated with the V-shaped groove formed in the surface of the semiconductor material of such devices. Figure 1 shows a cross-section of a VMOS transistor.

If the gate electrode is connected to the source and the drain contact at the bottom of Figure 1 is made positive relative to the source, no appreciable current will flow from drain to source, since the internal diode formed between the p and n type materials will be reverse biased. If, however, the gate electrode is made positive with respect to the source, the electric field produced by the gate potential creates a channel in the position shown in Figure 1. A current can now flow upwards from the drain through the channel to the source. As the gate becomes more positive, the width of the channel increases and the current from drain to source increases.
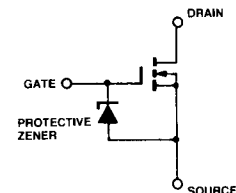

Figure 2. **Zener protection of the gate.**

If small changes in the gate voltage are to produce the required channel depth, the insulating layer must be extremely thin, which results in an appreciable gate input capacitance (typically some 50 pf). The thin layer also imposes a limit to the maximum voltage which can safely be applied to the gate without the risk of breaking down this thin layer and thus destroying the device. As the gate input resistance is so high (often of the order of a million megohms), it is very easy for small stray electrostatic charges to be picked up on the gate and produce voltages which can puncture the insulating film.

In some devices a small zener diode is connected between the gate and the source, as shown in Figure 2. If the gate to source voltage exceeds the zener voltage, the zener conducts and shorts out the voltage, protecting the MOSFET. However, the maximum zener current is quite small, so the zener can easily be damaged. The maximum input voltage
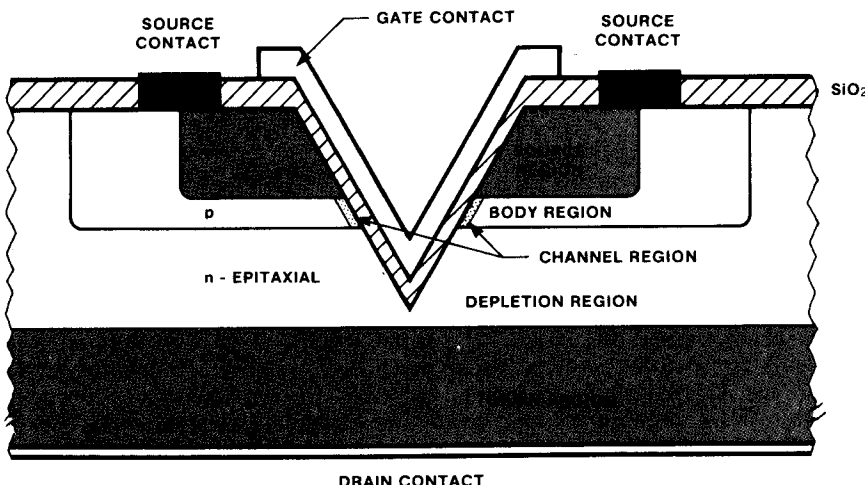

DRAIN CONTACT

Figure 1. **Structure of the VMOS device developed by Siliconix. With the gate biased positive with respect to the source, current flows from the drain region to the source via the channel region indicated. As the gate is biased more positive, the channel region increases, increasing the drain-source current. VMOS FETs are majority-carrier devices and can switch current in less than 10 ns. Bipolar transistors cannot compete as they suffer from minority carrier storage in the base region.**
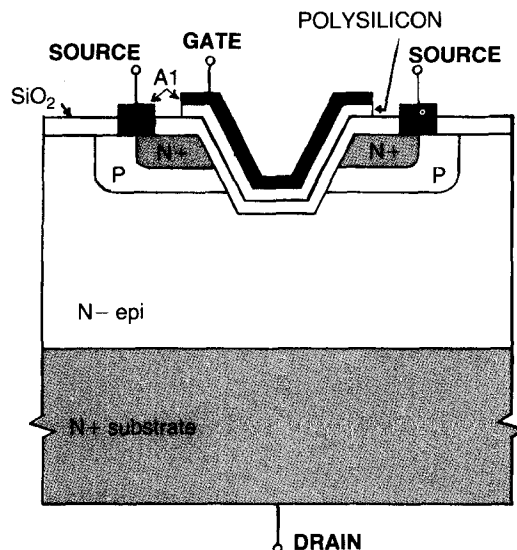
Figure 3. The U-groove device, introduced by Intersil, reduces problems associated with the intense electric field at the edge of the V-notch in VMOS devices. The polysilicon layer prevents migration of sodium impurity ions through the gate oxide layer, a source of chip failure in VMOS.
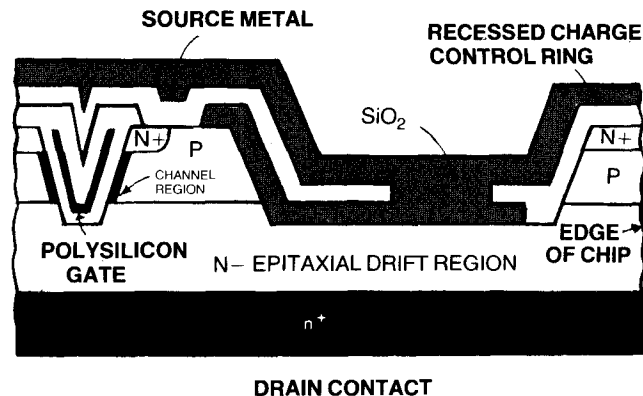

Figure 4. Having introduced power MOSFET technology, Siliconix have gone on to improve the devices. The tri-planar construction shown here allows much higher packing densities on the chip, the smaller size resulting in lower on state resistance. Polysilicon gates are buried in the oxide layers allowing source metallisation to cover a greater fraction of the chip area.

'in circuit' should not exceed the zener voltage so that the zener is used to provide protection against electrostatic charges only.

If the gate becomes more than a fraction of a volt negative with respect to the source, the zener will conduct in its forward direction. If one wishes to operate a MOSFET with the gate voltage negative with respect to the source at any part of the duty cycle, a device not containing a zener should be selected, but then one must take precautions to avoid electrostatic charge pick up.

The first VMOS devices marketed were n-channel devices, with an n-type channel formed in the p-type material shown in Figure 1. Siliconix still only market n-channel devices, but some other manufacturers offer both p-channel and n-channel types.

Comparison with bipolars

As the early VMOS devices could not handle so much current or so much applied voltage as conventional transistors, yet were more expensive than the latter, they obviously had some advantages or their manufacture would not have been a viable proposition.

Ordinary bipolar transistors suffer from the disadvantage of minority carrier storage in the base region. VMOS products are majority carrier devices and can therefore switch a current in less than 10 nanoseconds and operate up to several hundred mega-Hertz. For example, the 2N6657 can switch 1 A on or off in less than 4 ns, this being 10 to 200 times faster than a comparable bipolar device.

'Secondary breakdown' is another problem with bipolar transistors. If the current density increases at one point, the temperature rises in this region, leading to a still greater current density — a positive feedback effect which can lead to the rapid destruction of the device. In VMOS devices, an increase in the current density in the channel produces an increased temperature which results in a *lower* current density in that region, so that the current density automatically equalises itself throughout the chip without the formation of hot spots.

It follows that it is possible to connect two or more VMOS devices in parallel (often without any additional components), since the total current is automatically shared equally between the devices. Any device passing more current than the mean will become hotter and this will reduce the current somewhat in that device.

Apart from their higher cost, one of the disadvantages of VMOS devices is that their saturation voltage (typically 2V, maximum 4V for some devices when passing 1A) is much greater than for bipolar transistors. Although the V-shaped groove utilises the silicon area quite efficiently, the relatively sharp bottom of the groove is a disadvantage, since a strong electric field can be developed at this point between the gate and the drain where the insulating layer tends to be thinner than elsewhere. This results in a limited operating voltage capability owing to the possibility of gate to channel breakdown.

A perfect switching device would have an infinite resistance in the off state, but the drain current of many VMOS devices is in the nA region when in the off state with gate and source voltages equal. The resistance in the conducting state is normally a few ohms instead of the zero resistance of the perfect switch. This on-resistance is greater for devices with higher voltage ratings.

U-groove devices

The problem of the relatively intense electric field at the edge of the V-shaped notch of VMOS devices has already been mentioned. Intersil, followed by some other manufacturers, reduced this problem by producing devices with the structure shown in Figure 3, where the bottom of the groove is flat. Note that there is an additional layer of phosphorus-doped polycrystalline silicon between the gate and the insulating layer of silicon dioxide. This overcomes another problem of the early VMOS devices, namely the migration of sodium impurity ions through the gate oxide layer, which can cause reliability problems.

Other VMOS products

In 1980 Siliconix announced an improved triplanar VMOS process with the device structure shown in Figure 4. The source, the gate and the drain are each fabricated in a different plane. It is stated that this type of structure allows much higher packing densities on the chip and the smaller size will enable lower on state resistances to be obtained. Polysilicon gates are buried under the oxide layers so that the source metallisation can cover a greater fraction of the chip area.

Another major improvement from the triplanar structure arises from the use▶

of thin low-resistivity doped layers and from a re-arrangement of the V grooves for optimum use of the epitaxial layers.

Vertical DMOS

Although the modified VMOS processes are very good for devices rated up to about 150 V, they are not ideal for higher voltages. The vertical DMOS structure shown in Figure 5 has been found very suitable for high voltage devices. The current flows upwards from the drain into the n-epitaxial layer, but then flows horizontally for a short distance through a channel to the source.

Supertex of California use this technique to make devices with ratings of up to about 500 V, but somewhat higher voltage devices of this type are likely to become available. Figure 5 shows how the main junction region is surrounded by a concentric second junction which is in turn surrounded by a third junction. Apart from high voltage capability, this process can produce devices with a very low on-resistance (down to 0.05 ohm). In addition the devices are very fast, owing to the low gate capacitance. For example, a 1A device can operate at about 2 GHz and a 10 A device at about 500 MHz.

The Ferranti Company of Oldham, England have co-operated with Supertex to develop vertical DMOS devices, both n-channel and p-channel, with ratings up to 650 V and drain currents up to 16 A continuous.
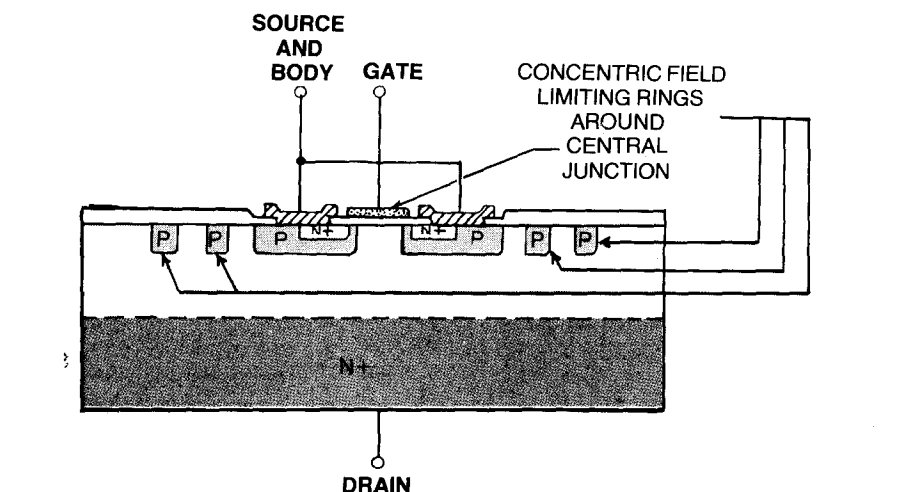
Figure 5. In the vertical DMOS device, current flows from the drain (N+) into the n-epitaxial layer (N−) then flows horizontally through a channel into the source. The concentric rings of p-type material around the main junction help improve the current capability and reduce the on-resistance. This form of construction achieves significantly higher voltage and current ratings compared to prior power MOSFETs.

Hitachi devices

Hitachi has developed a MOSFET device with the structure shown in Figure 6. The gate oxide layer is designed to handle only 20 to 30 V, so a field plate is provided to prevent high electric fields from forming near the gate. This type of device is most suitable for audio frequencies and for operation at up to a maximum of a few MHz. Both p-channel and n-channel types are available with ratings of up to 200 V and 8 A.
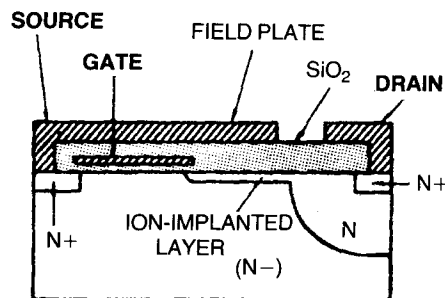
Figure 6. Hitachi MOSFET construction.

HEXFET devices

In mid-1980 International Rectifier introduced a range of devices named HEXFET after the hexagonal structure of the source cells which are connected by a common silicon gate (see Figure 7). The density of these source cells is over half a million per square inch.

HEXFET devices are available in both p-channel and n-channel polarities and can handle high power levels. They have voltage ratings of up to 500 V and continuous current ratings of up to 25 A. Values of channel resistance as low as 0.05 ohm can be obtained in the on state.

Some of the main applications for HEXFETS include servo motor control, RF induction heating, welding control equipment, audio amplification and other uses where the control of high power is required.

SIPMOS

The latest technology to emerge in the power MOSFET field is SIPMOS from Siemens of West Germany, which is an extension of the vertical DMOS technique. Siemens have used this technique to fabricate the first 1000 V MOSFET device, the BUZ 54, which can handle 5 A. It is expected to have wide ▶
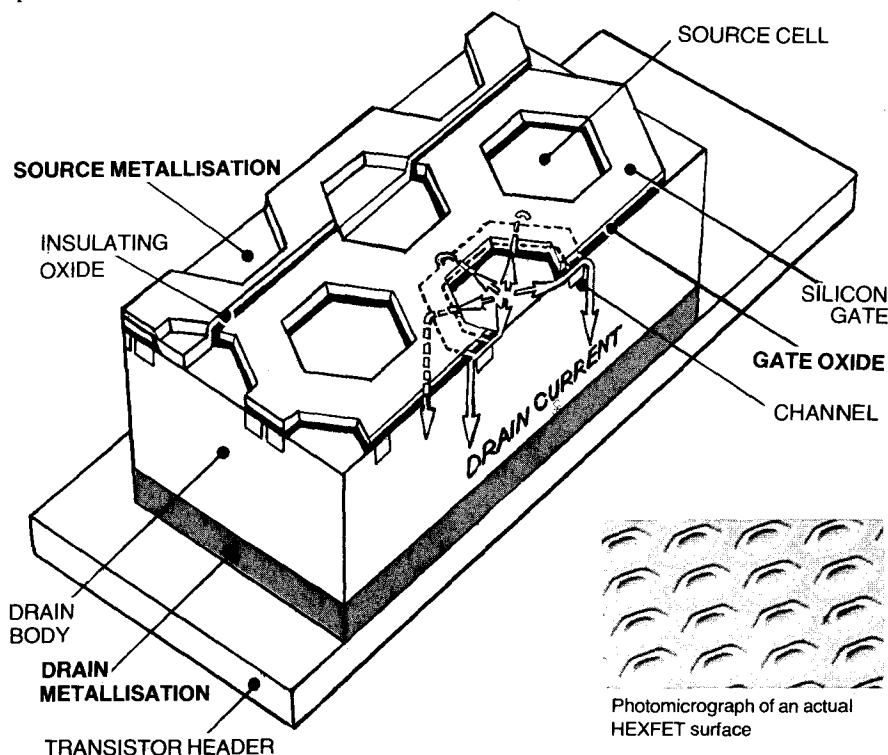

Photomicrograph of an actual HEXFET surface

Figure 7. **Construction of the HEXFET device introduced by the International Rectifier company in mid-1980. The hexagonal source cells (hence the name) are connected by a common silicon gate. Claimed advantages include high voltage and current ratings plus very low on-resistance.**

The saturation characteristics of a VMOS device (left) and a HEXFET (right)
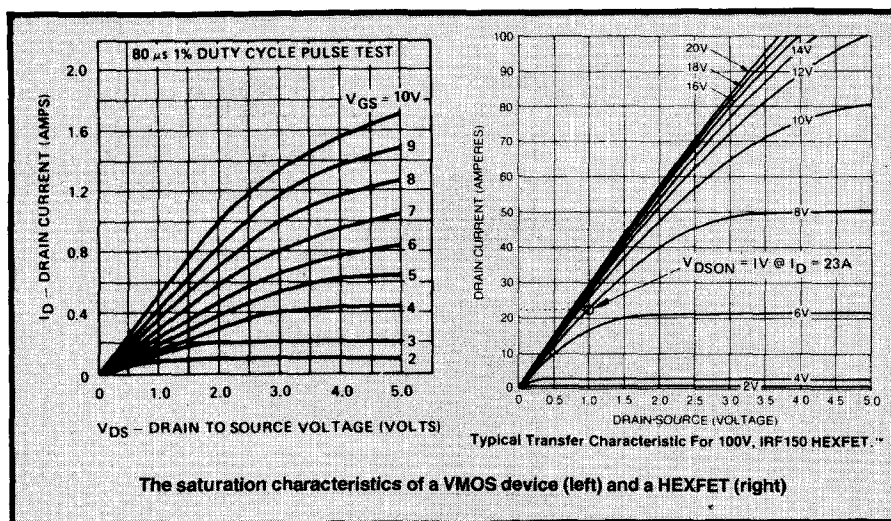
uses in switching mode power supplies.

Other SIPMOS devices have ratings in the range of 50 V to 500 V, all being n-channel types. SIPMOS transistors can switch loads of up to 5 kW using inputs to the gate of less than 1 mA at 5 V. Maximum drain currents of up to 30 A can be handled, while on-resistance values can be as low as 0.03 ohm.

Applications

Power MOSFET devices can be used as alternatives to power transistors and power Darlington devices in many applications, but they are generally more expensive than the latter and the circuit designer must decide which types of device are most suitable for his own application.

The use of power MOSFET products is particularly attractive when one can take advantage of their high switching speed or their high frequency capability. Although they may be somewhat more expensive than other transistors, the use of these new devices may simplify circuitry and reduce the overall costs. For example, a conventional power transistor requires a considerable current at its input and one or more driver stages may be required to provide this current, whereas the high input impedance of the power MOSFET enables the latter to operate with such small input currents that power driver stages can usually be eliminated.
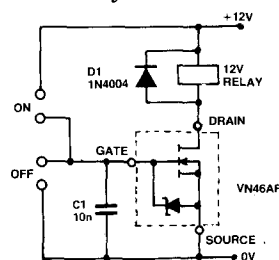
Figure 8. Simple touch switch.

Simple touch switch

The circuit of Figure 8 shows how the very high input impedance of a VMOS power MOSFET can be employed in a simple touch switch. When the circuit is first switched on, the capacitor C1 is normally fully discharged, so the VN46AF VMOS device passes negligible drain current.

When the upper pair of contacts is touched, current flows from the +12 V line ,through the person's skin and charges C1. The VN46AF device is thus biased to conduction and the relay closes. If a finger is now placed across the lower touch contacts, C1 discharges and the VN46AF is turned off, opening the relay. The diode D1 is used to bypass the transient voltages formed when the current ceases to flow through the relay coil — such voltages can destroy MOSFETs.
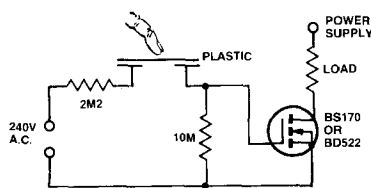
Figure 9. Capacitive touch switch.

Capacitive touch switch

The gate circuit impedance of VMOS devices is so high that circuits can be designed as touch switches in which no part of the circuit is actually touched. In Figure 9 (designed by ITT Semiconductors), the presence of a finger just above the plastic material at the point of separation of the electrodes under the plastic is sufficient to cause current to flow in the load.

The capacitance between each of the electrodes and the finger allows a small alternating current to flow through the 2M2 safety resistor to the gate circuit of the small BS170 or the larger BD522 n-channel device.
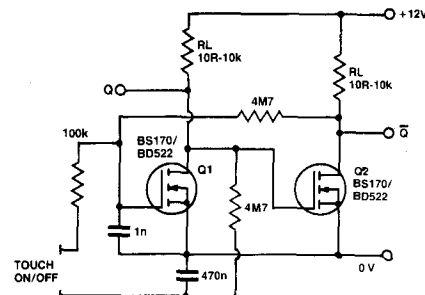

Figure 10. Capacitive touch switch will cycle on and off if finger is held on the sensor.

Figure 10 shows another touch switch designed by ITT Semiconductors, only a single touch point being used for on/off operation. When power is first switched on, T1 will conduct and T2 is kept non-conducting. Touching the sensor contacts will cause T2 to conduct and feedback from the drain of this device through the 4M7 resistor to the gate of T1 will keep the latter device in the non-conducting state. The 470n capacitor now becomes charged.

If the sensor is touched again, the positive potential from this capacitor is transferred to the gate of T1 and the latter device is switched to conduction, whilst T2 is turned off. If the sensor is touched for longer than about one second, the circuit will operate as a relaxation oscillator which changes its state about once per second. The load impedances employed in this circuit need not be identical, any values from about 10 ohm to 10k being suitable.

CMOS Interfacing

The 4000 series of CMOS logic devices can provide only small output currents, but sometimes one wishes to use the output from such a device to control a relay or other load which requires a relatively large current. A VMOS device can conveniently be employed to match the high output impedance of a CMOS device to a relatively low load
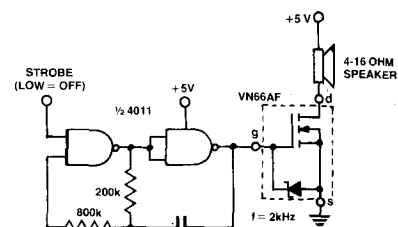

Figure 11. Audio alarm.

impedance such as a tungsten filament lamp.

An example is the audio alarm circuit of Figure 11. Two of the four logic gates of a CD4011 device are connected as a standard 2 kHz oscillator. Any appreciable current taken from the output of this oscillator affects the operation of the circuit, but the VN66AF requires negligible current and forms an ideal

interface bevice between the CMOS oscillator and the loudspeaker.

When the upper input of the left hand CMOS gate is connected to the +5 V line, oscillation takes place, but when this input is connected to ground, oscillation ceases. Thus a high impedance logic output can be used to switch the oscillator on and off through the use of this input to the left hand gate.

Figure 12 is an interesting variation of the circuit of Figure 11 in which the four gates of a 4011 device are used to form two oscillators. The two left hand gates form a sub-audio frequency oscillator which modulates the audio oscillator formed by the two right hand gates of Figure 12. Thus one obtains a much more impressive two-tone alarm sound than with the simpler constant-note circuit of Figure 11.

The timer circuit of Figure 13 is another example of VMOS interfacing between a CMOS device and a relay. In the quiescent state, the upper input to the left hand gate will be low and the output from this gate high. Thus the output from the right hand gate will be low and the relay will remain open.

If the start switch is momentarily closed, the high input applied to one input of the left hand gate will cause the output from this gate to go low, while the output from the right hand gate goes high and switches the VN46AF to conduction. Thus the relay closes.

The capacitor between the two gates charges slowly through the fixed and variable resistor from the positive supply line. When the inputs to the right hand gate become sufficiently high in potential, the output of this gate goes low and by feedback to the left hand gate the circuit switches back rapidly to its quiescent state in which
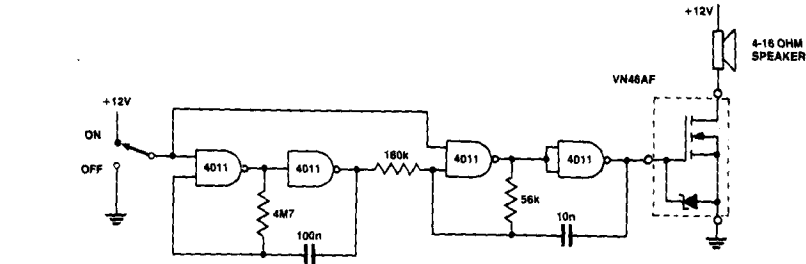
Figure 12. This two-tone alarm is a variation of the Figure 11 circuit.
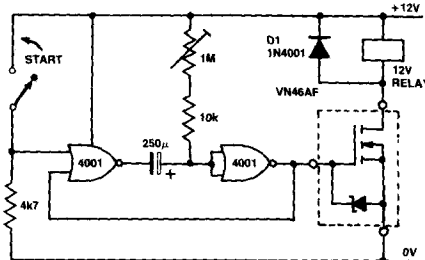
Figure 13. Simple timer has a variable range from a few seconds to a few minutes. The 1M pot sets the time the relay holds in.

Figure 14. Simple delay timer. The load could be a lamp, relay or whatever.

negligible current passes through the relay. The length of time for which the relay remains closed can be set by the 1M pot or by altering the value of the capacitor connected between the two gates. When the values shown are used, times obtained range from a few seconds to a few minutes as the variable resistor is moved.

Delay switch

A simple VMOS delay switch is shown in Figure 14. When the switch is closed for a moment, the capacitor becomes fully charged and the VN46AF passes current through the load. The capacitor slowly discharges through the 10M resistor, so the gate voltage of the VN46AF will eventually fall to a value where very little current can pass through the load.
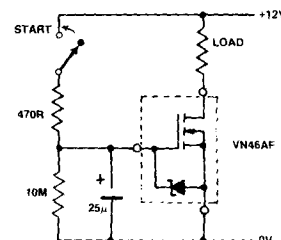
Auto applications

The fast switching ability of MOSFET devices renders them very suitable for use in vehicle electronic ignition systems. Timing pulses from a magnetic or other contactless pickup may be fed to an IC which provides a voltage output for the control of a MOSFET device. The latter switches the current through an ignition coil to provide the required high voltage.

An automobile circuit using a SIPMOS transistor as a power switch is shown in Figure 15. As in so many applications of MOSFET devices, the high input impedance of the SIPMOS device is utilised here, since it can be voltage driven by a suitable IC. This circuit is for an automobile alternator voltage regulator and has been designed for a SIPMOS device rated at ▶
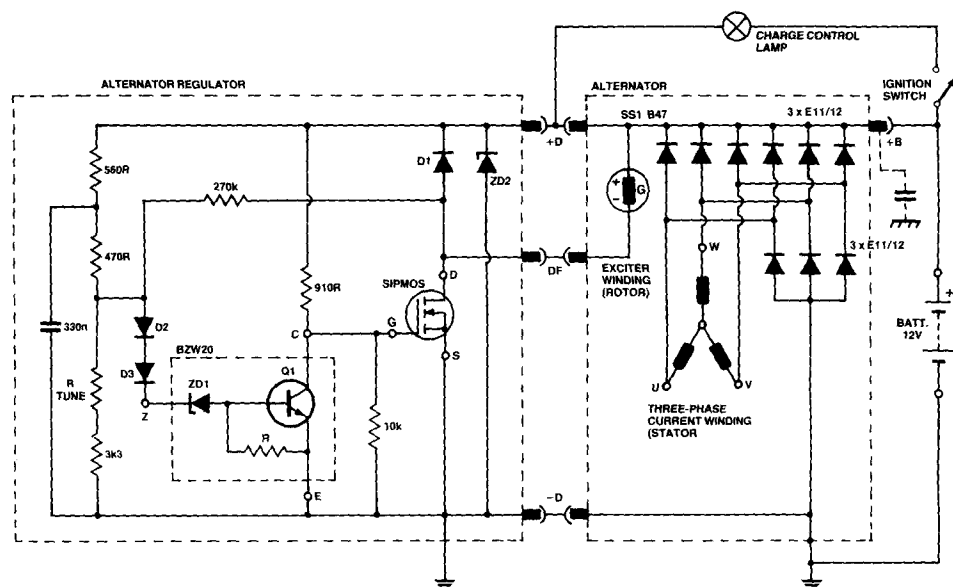

Figure 15. The latest power MOSFET development, SIPMOS, has already found application in automotive electronics. This circuit is an alternator regulator and employs a SIPMOS device rated at 500 V/8 A and an on-resistance not greater than 0.2 ohm.
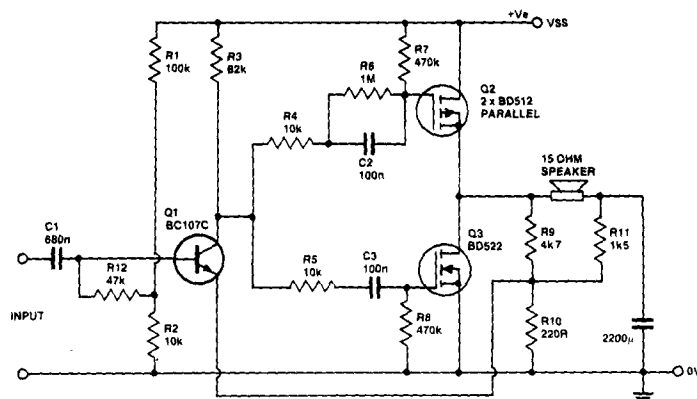
Figure 17. **Class ABC amplifier circuit from ITT Semiconductors is simple but has 3½% distortion at 1.75 W and is only suited to general applications.**

about 500 V maximum drain-to-source voltage, 8 A current and an on-resistance of not more than 0.2 ohm.

Simple audio applications

The excellent linearity of VMOS devices has attracted considerable interest in their possible use in the audio field but the relatively high price of these devices and their previously limited power handling capability retarded their adoption until recently. They may be used in simple, low-power circuits, but moderately high power ultra-low distortion circuits have also been designed using VMOS devices. The very fast switching ability of VMOS devices also makes them very suitable for Class D pulse width modulation circuits.

To operate a VMOS device as a simple class A amplifier, it is only necessary to provide a bias network so that the device operates in its linear region without cutoff. The gate is connected to a tap on a resistive potential divider across the power supply lines and the input signal is capacitively coupled to the gate. The gain will be approximately equal to the mutual conductance of the device multiplied by the load resistance; gain values of over 30 dB are obtainable, and this gain extends well into the MHz region.

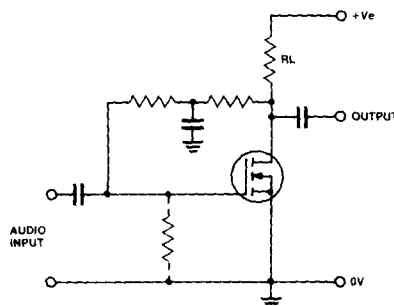A circuit of this general type is shown in Figure 16. The bias level of the

Figure 16. **General circuit of a simple class-A audio amplifier using a power MOSFET.**

VMOS device is stabilised by means of negative voltage feedback from the drain to the gate circuit.

Figure 17 shows a particularly interesting circuit from ITT Semiconductors which they call a class ABC amplifier, since it is basically a Class B amplifier, but one of the transistors is more in Class A, while the other is definitely in Class C. It is a simple circuit not designed for particularly low distortion.

The output stage is unusual in that it comprises two BD512 p-channel VMOS devices in one part and a complementary BD522 single device in the other part. This is because hole mobility in the p-channel BD512 is only half that of the electron mobility in the n-channel BD522s so two p-channel devices are required to obtain about the same mutual conductance as that provided by the single n-channel device. As explained earlier, MOSFETs can be connected in parallel without extra circuitry because they automatically share the current.

As the drain electrodes of the VMOS devices in Figure 17 are connected to the device tabs, all of the tabs can be bolted to the same heatsink without the need for insulating washers. The negative feedback circuit compensates for any variations in the biasing requirements of the particular VMOS devices employed. Both ac and dc feedback are employed, but there is heavier dc feedback through R11 and R10 to stabilise the quiescent dc output voltage at half the supply potential so as to ensure a maximum available output voltage swing.

This circuit provides a voltage gain of 30 and a bandwidth extending from 35 Hz to 125 kHz at the −6 dB points. Distortion increases at ultrasonic frequencies above about 25 kHz (as with most audio amplifiers). When a 25 V supply is used, the distortion is a minimum of about 0.4% at about 0.5 W, rising to about 0.8% at 1 W, 2% at 1.5 W and 3½% at 1.75 W.

High fidelity

In 1976 Siliconix published a circuit for a high quality 40 W amplifier using VMOS devices, but each half of the output stage required three VMP12 (now designated 2N6658), 90 V TO-3 devices in parallel. Thus, twelve of the devices were required in a stereo amplifier providing 40 W per channel. Rather cumbersome — and costly. However, distortion at the mid-frequency range was only about 0.04% at the 40 W level and about 0.025% at the 1 W level. Only 22 dB of feedback was needed to obtain a response flat to 4 MHz and the slew rate was 100V/μs!

One of the advantages claimed for VMOS amplifiers is the lack of transient intermodulation (TIM), because the power bandwidth exceeds the small signal bandwidth. For any frequency below 500 kHz, the amplifier simply overloads before TIM appears.

Taking things a step further, the circuit in Figure 18 is a simple power amplifier first published in the Hitachi MOSFET application notes. The 2SK133 and 2SJ48 have an on-resistance of roughly two ohms, so that at 7 A peak output current you can expect a voltage drop of about 14 V across each device. With the power supply voltages shown the circuit is capable of around 50 W.

Transistors Q1 and Q2 form an input differential pair that compares the input signal with the output signal of the amplifier. The difference between these two signal voltages is fed to a second differential pair, Q4 and Q5. This ensures that the open-loop voltage gain of the amplifier is high and allows a fairly high feedback factor when negative feedback is applied. A relatively large amount of negative feedback is essential when using MOSFETs like this in audio amplifiers to linearise the MOSFET characteristics which have, on average, *10 times* the distortion of a typical bipolar transistor of similar
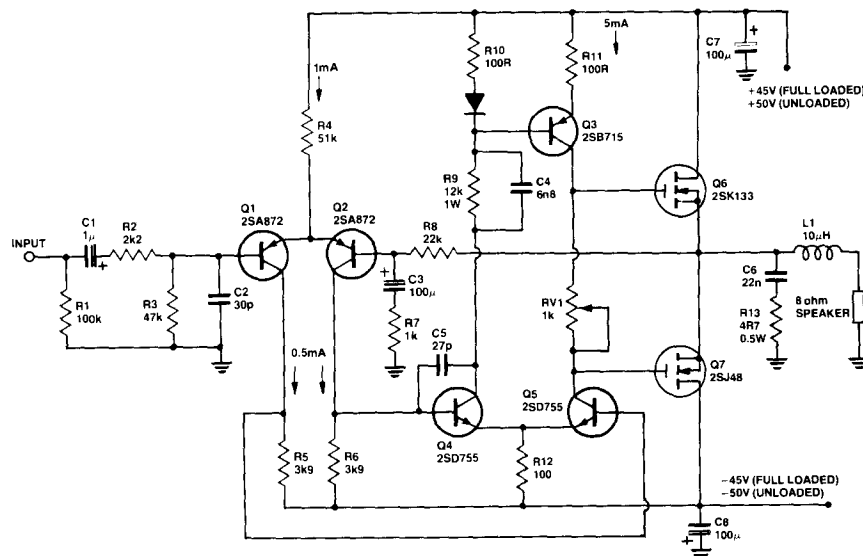

Figure 18. **Circuit of a 50 W hi-fi amplifier from the Hitachi MOSFET application notes. Performance is quite good but dependent on the driver transistors.**

power capabilities.

The transistors forming the driver stage, Q3 and Q5, have been specially designed by Hitachi to drive MOSFETs. They're superb devices, having a V_{ceo} of 100 V and a typical gain (h_{FE}) of around 500. With these transistors the distortion characteristics shown in Figure 19 can be expected. Unfortunately, these transistors are not available in Australia at the present time and substituting alternative available transistors degrades performance considerably. A BD139/BD140 complementary pair for instance, with typical h_{FE} of around 50, is not capable of providing the necessary open-loop gain, especially at high frequencies. An experimental circuit we built with BC177s and BD139/BD140s gave less than 0.02% at 1 kHz at full power, rising to as much as 0.1% or more at 20 kHz. So, MOSFETs with all of their advantages have disadvantages too — mainly due to the fact that the forward transconductance is only a fraction of that of a good bipolar transistor.

In order to design an extremely high quality amplifier employing MOSFETs, we are really faced with a new set of problems to solve, but with the promise of performance that makes it worthwhile. (This section on Figure 18 inserted by David Tilbrook . . . Ed.). ●
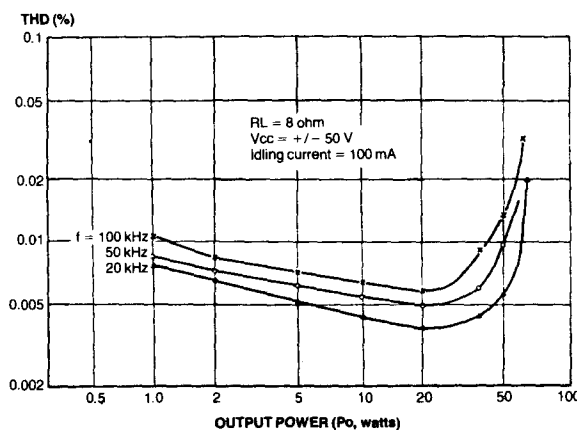

Figure 19. **Distortion characteristics of the circuit in Figure 18.**

Designing potcore inductors

Potcores offer many advantages when you need to use an inductor in a circuit. This article details a simplified design method for two common potcore sizes based on a nomograph.

FERRITE POTCORES are widely used in the construction of small inductors and transformers. However very few amateurs know how to choose a core appropriate to their needs, or how to wind a coil of specific inductance.

This article describes a simple method of designing coils for low frequency applications. The design of coils for high frequency applications, and of transformers, is beyond the scope of this article. Design details given apply only to the Philips 'P' series of pot cores and more particularly to the 18 mm (P18/11) and 26 mm (P26/16) diameter cores which are the most commonly available.

Each core size is available in four different ferrite materials (3H1, 3B7, 3D3, 4C6) to cover the frequency range from audio to about 20 MHz. Additionally each material, in each size, is available with a number of permeabilities to cover different inductance, stability and Q factor requirements.

There are two factors commonly used to classify ferrite cores. These are effective permeability (μe) and A_L factor.

The μe factor is primarily determined by the permeability of the material used and its cross sectional area, and secondly by the air gap left between the centres of the two core halves. For example an 18 mm 3B7 core without any air gap (type 0 4000) has a μe of 1750 and a tolerance on inductance of $\pm 25\%$. The use of increasingly larger air gaps in the same core size and material lowers the μe but increases the stability and reduces the tolerance on inductance.

A second factor in common usage is A_L. This factor gives, in nanohenries, the inductance of ONE turn on the core. The inductance of N turns on the core is

$L = N^2 A_L \times 10^{-3}$ millihenries

The selection of a core size, a core material, a permeability value, a wire size and the number of turns depends on all the following factors:-

— inductance, stability of inductance
— frequency range
— Q factor
— unbalanced dc coil current
— level of ac coil current

Choosing the correct core taking *all* these factors into account is a difficult task indeed. However a large number of
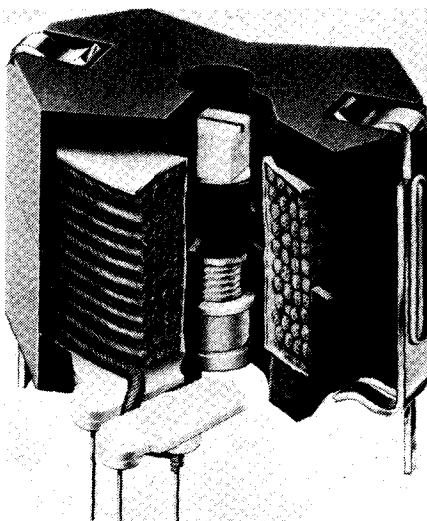
Cutaway view of a potcore showing the winding on the internal bobbin and the inductance adjustor in the centre. The ferrite 'pot' comes in two halves, held together with a clip.

core types are eliminated by first selecting in accordance with frequency range and stability.

Frequency range

Firstly select the core material from Table 1 in accordance with the desired frequency range. To choose between 3H1 and 3B7 it is necessary to consider temperature stability.

If the tuning capacitors associated with the coil have small or varying temperature coefficients, a 3B7 core should be used as they have the lowest temperature coefficient in the range 0° - 70°C. Alternatively, if using polystrene capacitors (temp. coeff. -150 ppm/°C) a 3H1 core having an effective permeability (μe) around 150 will give excellent temperature compensation for the temperature coefficient of these capacitors.

Inductance stability

Since the inductance of a coil is

TABLE 1	
FREQUENCY RANGE	**CORE TYPE**
0.1 — 200 kHz	3B7, 3H1
200 kHz — 2 MHz	3D3
2 MHz — 20 MHz	4C6

proportional to core permeability, the change of effective permeability (μe) with temperature determines the stability of inductance.

The percentage change of inductance with temperature is linearly proportional to μe and hence low μe cores should be used for greatest stability. Stability is therefore obtained at the expense of inductance obtainable with a given core size.

The temperature effect is not large enough to affect any but the most critical of applications and the tolerance on inductance as stated in Tables 2 and 3 will be obtained over the temperature range $+15°$ to $+35°$C.

Direct current

A direct current in the winding will change the inductance value of the core and if large enough, could cause saturation. In general, large air gaps, and hence lower permeability (μe), cores should be used where large dc currents are flowing.
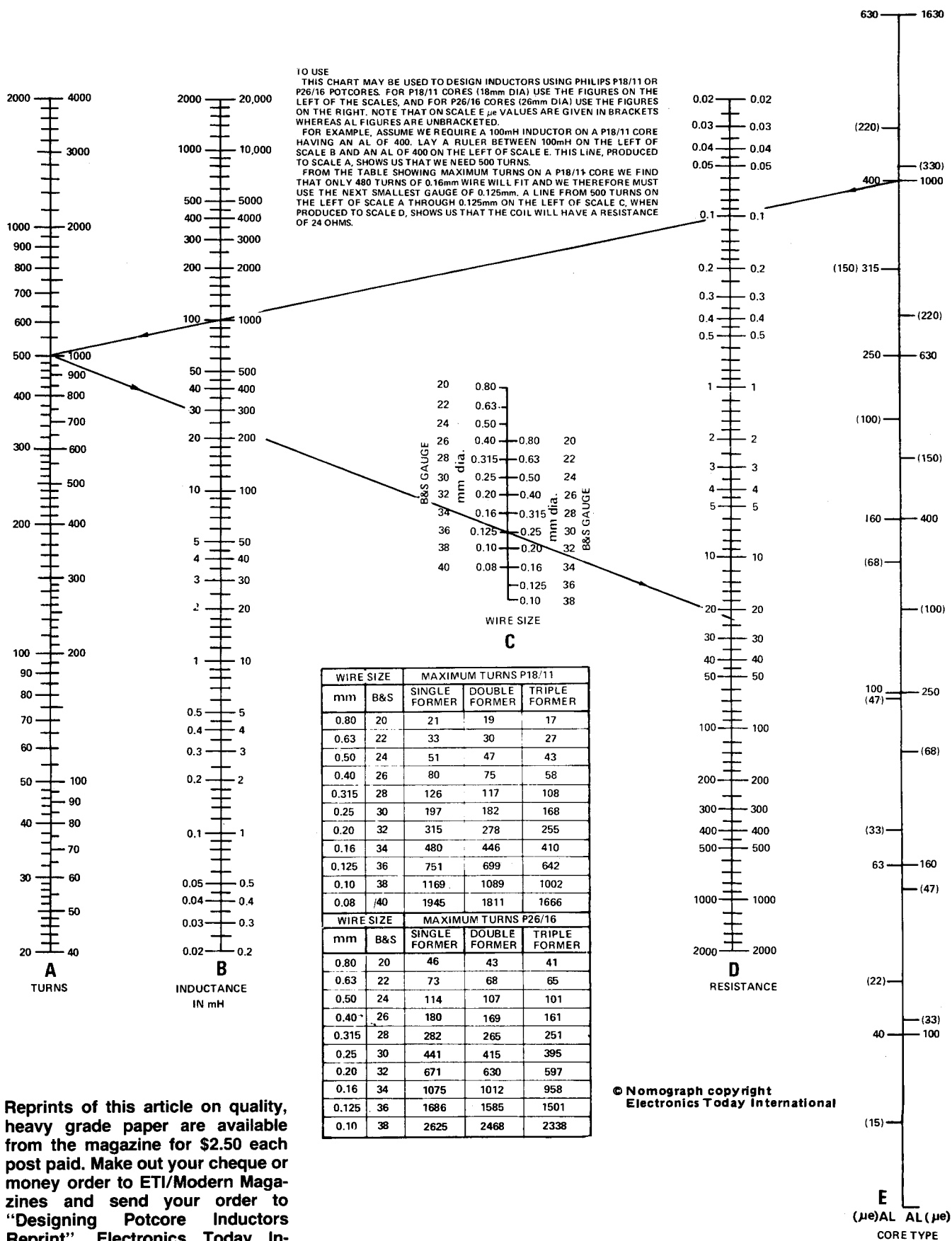
Q factor

The Q of a coil is influenced by different factors at different frequenceis.

At frequencies below 10kHz it is almost completely determined by the dc resistance of the winding. The Q factor of any given coil increases linearly with frequency, and the larger the core, the larger the Q. The highest Q factors are obtainable by using gapless cores of 3H1 or 3B7 material (providing that tolerance and stability are acceptable) eg. 04000 series (P18/11) and 08000 series (P26/16).

Throughout the ultrasonic range core and winding losses affect Q, but Q factors of several hundred may still be obtained by optimum choice of wire and core, such that core and winding losses are equal. For further information, on optimum design, reference should be made to Philips Data Handbook — Components and Materials, Vol. 4.

At higher ultrasonic and lower radio frequencies, additional factors of dielectric and skin-effect losses and parallel winding capacitance, all affect Q, making exact design difficult. Use of Litz wires, split section formers and small cores with low μe values will assist.

IO USE

THIS CHART MAY BE USED TO DESIGN INDUCTORS USING PHILIPS P18/11 OR P26/16 POTCORES. FOR P18/11 CORES (18mm DIA) USE THE FIGURES ON THE LEFT OF THE SCALES, AND FOR P26/16 CORES (26mm DIA) USE THE FIGURES ON THE RIGHT. NOTE THAT ON SCALE E μe VALUES ARE GIVEN IN BRACKETS WHEREAS AL FIGURES ARE UNBRACKETED.

FOR EXAMPLE, ASSUME WE REQUIRE A 100mH INDUCTOR ON A P18/11 CORE HAVING AN AL OF 400. LAY A RULER BETWEEN 100mH ON THE LEFT OF SCALE B AND AN AL OF 400 ON THE LEFT OF SCALE E. THIS LINE, PRODUCED TO SCALE A, SHOWS US THAT WE NEED 500 TURNS.

FROM THE TABLE SHOWING MAXIMUM TURNS ON A P18/11 CORE WE FIND THAT ONLY 480 TURNS OF 0.16mm WIRE WILL FIT AND WE THEREFORE MUST USE THE NEXT SMALLEST GAUGE OF 0.125mm. A LINE FROM 500 TURNS ON THE LEFT OF SCALE A THROUGH 0.125mm ON THE LEFT OF SCALE C, WHEN PRODUCED TO SCALE D, SHOWS US THAT THE COIL WILL HAVE A RESISTANCE OF 24 OHMS.

A — TURNS

B — INDUCTANCE IN mH

C — WIRE SIZE

D — RESISTANCE

E — (μe)AL AL (μe) CORE TYPE

WIRE SIZE		MAXIMUM TURNS P18/11		
mm	B&S	SINGLE FORMER	DOUBLE FORMER	TRIPLE FORMER
0.80	20	21	19	17
0.63	22	33	30	27
0.50	24	51	47	43
0.40	26	80	75	58
0.315	28	126	117	108
0.25	30	197	182	168
0.20	32	315	278	255
0.16	34	480	446	410
0.125	36	751	699	642
0.10	38	1169	1089	1002
0.08	/40	1945	1811	1666

WIRE SIZE		MAXIMUM TURNS P26/16		
mm	B&S	SINGLE FORMER	DOUBLE FORMER	TRIPLE FORMER
0.80	20	46	43	41
0.63	22	73	68	65
0.50	24	114	107	101
0.40	26	180	169	161
0.315	28	282	265	251
0.25	30	441	415	395
0.20	32	671	630	597
0.16	34	1075	1012	958
0.125	36	1686	1585	1501
0.10	38	2625	2468	2338

© Nomograph copyright
Electronics Today International

Table 2 P18/11 Potcores

A. PRE-ADJUSTED PAIRS WITH STANDARD μ_e VALUES

catalogue number 4322.022 ...	grade of ferroxcube	effective permeability (μ_e)	number of turns for 1 mH α	tolerance on inductance %	adjustor type 4322.021	adjustor colour
28030	3B7	33	98.2	±1	30780	green
28040	3B7	47	82.3	±1	30800	red
28050	3B7	68	68.4	±1	30980	white
28060	3B7	100	56.4	±1.5	30980	white
28070	3B7	150	46.1	±2	30810	brown
28080	3B7	220	38.1	±3	30810	brown
28090	3B7	330	31.0	±3	31090	grey
28230	3H1	33	98.2	±1	30780	green
28240	3H1	47	82.3	±1	30800	red
28250	3H1	68	68.4	±1	30980	white
28260	3H1	100	56.4	±1.5	30980	white
28270	3H1	150	46.1	±2	30810	brown
28280	3H1	220	38.1	±3	30810	brown
28290	3H1	330	31.0	±3	31090	grey
28430	3D3	33	98.2	±1	30780	green
28440	3D3	47	82.3	±1	30800	red
28450	3D3	68	68.4	±1	30980	white
28810	4C6	15	146	±1	30780	green
28830	4C6	33	98.2	±1	30790	yellow
08000	3B7	1910.0	12.9	±25	–	–
08200	3H1	1910.0	12.9	±25	–	–
08400	3D3	730.0	20.8	±25	–	–

$$N = \alpha\sqrt{L} \ (L \ in \ 10^{-3} \ H)$$

B. PRE-ADJUSTED PAIRS WITH STANDARD AL FACTORS

catalogue number 4322.022 ...	grade of ferroxcube	AL factor	corres- ponding μ_e value	tolerance on inductance %	adjustor type 4322.021	adjustor colour
29030	3B7	63	20	±1	30780	green
29040	3B6	100	31.8	±1	30780	green
29050	3B7	160	51	±1	30800	red
29060	3B7	250	79.5	±1	30980	white
29070	3B7	315	100.2	±1.5	30980	white
29080	3B7	400	127	±2	30810	brown
29100	3B7	630	200	±3	30810	brown
29110	3B7	1000	318	±3	31090	grey
29120	3B7	1600	510	±3	31090	grey
29230	3H1	63	20	±1	30780	green
29240	3H1	100	31.8	±1	30780	green
29280	3H1	160	51	±1	30800	red
29260	3H1	250	79.5	±1	30980	white
29270	3H1	315	100.2	±1.5	30980	white
29280	3H1	400	127	±2	30810	brown
29300	3H1	630	200	±3	30810	brown
29310	3H1	1000	318	±3	31090	grey
29320	3H1	1600	510	±3	31090	grey
29430	3D3	63	20	±1	30780	green
29440	3D3	100	31.8	±1	30780	green
29450	3D3	160	51	±1	30800	red
29460	3D3	250	79.5	±1	30980	white
29830	4C6	63	20	±1	30780	green
29840	4C6	100	31.8	±1	30790	yellow

$$L = N^2 \ AL \ (10^{-9} \ H)$$

C. COILFORMERS

catalogue number	number of sections
4322.021.30330	1
4322.021.30340	2
4322.021.30350	3
4322.021.30130	1 with pins
4302.021.20030	1 with pins

D. MOUNTING PARTS

catalogue number	description
4322.021.30550	container
4322.021.30660	spring
4322.021.30470	tag plate
4322.021.30710	nut
4322.021.30720	bush
4302.021.20020	clip

Inductance

The tolerance given on inductance, in Tables 2 and 3 is obtained when using the specified core, and a wire size that will completely fill the former close layer wound. Due to slight changes in wire diameter and different methods of winding, the exact number of turns accommodated may vary by ±10%.

Table 3 P26/16 Potcores

A. PRE-ADJUSTED PAIRS WITH STANDARD μ_e VALUES

catalogue number 4322.022	grade of ferroxcube	effective permeability (μ_e)	number of turns $\propto$	tolerance on inductance %	adjustor type number 4322.021 ...	adjustor colour
24030	387	33	120	±1	30760	green
24040	387	47	100.5	±1	30770	red
24050	387	68	83.6	±1	30960	yellow
24060	387	100	68.9	±1.5	30970	white
24070	387	150	56.3	±2	30730	brown
24080	387	220	46.5	±3	31080	grey
24230	3H1	33	120	±1	30760	green
24240	3H1	100.5	100.5	±1	30770	red
24250	3H1	68	83.6	±1	30960	yellow
24260	3H1	100	68.9	±1.5	30970	white
24270	3H1	150	56.3	±2	30730	brown
24280	3H1	220	46.5	±3	31080	grey
24430	3D3	33	120	±1	30760	green
24440	3D3	47	100.5	±1	30770	red
24450	3D3	68	83.6	±3	30960	yellow
24810	4C6	15	178	±1	30760	green
24820	4C6	22	147	±1	30770	red
24830	4C6	33	120	±1	30970	white
04000	3B7	1750	16.5	±25	—	—
04200	3H1	1750	16.5	±25		
04000	3D3	705	25.9	±25		

$$N = \propto \sqrt{L}\ (L\ \text{in}\ 10^{-3}\ H)$$

B. PRE-ADJUSTED PAIRS WITH STANDARD AL FACTORS

catalogue number 4322.022	grade of ferroxcube	AL factor	corresponding μ_e value	tolerance on inductance $\propto$	adjustor type 4322.021 ...	adjustor colour
25030	3B7	63	30	±1	30760	green
25040	3B7	100	47.5	±1	30770	red
25050	3B7	160	76	±1	30960	yellow
25050	3B7	250	119	±1.5	30970	white
25070	3B7	315	149	±2	30730	brown
25080	3B7	400	190	±2	31080	grey
25100	3B7	630	298	±3	31080	grey
25230	3H1	63	30	±1	30760	green
25240	3H1	100	47.5	±1	30770	red
25250	3H1	160	76	±1	30960	yellow
25260	3H1	250	119	±1.5	30970	white
25270	3H1	315	149	±2	30730	brown
25280	3H1	400	190	±2	31080	grey
25300	3H1	630	298	±3	31080	grey
25420	3D3	40	19.0	±1	30760	green
25430	3D3	63	30	±1	30760	green
25440	3D3	100	47.5	±1	30770	red
25450	3D3	160	76	±1	30960	yellow
25810	4C6	25	11.9	±1	30760	green
25820	4C6	40	19.0	±1	30770	red
25830	4C6	63	30	±1	30970	white

Hence it is safer, when winding experimental coils, to only try and fit 90% of the turns indicated in the maximum number of turns tables.

If the former is only partly filled, errors up to 4% may occur with the lower μe cores. However the use of an adjustor will allow a +10% adjustment range which is generally sufficient to cope with tolerances found in practical circuits.

When optimum stability is required the type of adjustor that matches a certain core should be used. If it is desired to widen the adjustable range, at the possible expense of stability, an adjustor indicated for a potcore with a high μe value may be used with a potcore of low μe value.

C. COIL FORMERS

catalogue number	number of sections
4322.02130270	1
4322.021.30280	2
4322.021.30290	3
4322.021.30090	1 with pins
4302.021.20010	1 with pins

C. MOUNTING PARTS

catalogue number	description
4322.021.30530	container
4322.021.30640	spring
4322.021.30450	tag plate
4322.021.30710	nut
4322.021.30720	bush
4302.021.20000	clip

Don't go off about Schmitt triggers — look at the 4093!

This highly versatile package can be used in a wide variety of applications, including: wave-shaping, timing, logic circuits and waveform generation. Ray Marston explains.

ONE OF THE HANDIEST 'building blocks' in electronic circuitry is the Schmitt trigger — a simple circuit block whose output changes state when the input goes above or below a certain "threshold". The rising input threshold may be set at one level and the falling input threshold may be set at another level — usually below the former. The difference between the two levels is called the threshold "hysteresis".

The wonders of small-scale integration can now bring you four Schmitt triggers in a single package! What we are discussing this month is the 4093 CMOS IC.

This device is a quad two-input NAND Schmitt trigger — to use the jargon. It is a highly versatile package that can be used in a wide variety of applications, including: wave-shaping, timing, logic circuits and waveform generation.

Schmitt Applications

Figure 1 shows the functional diagram and truth table of the 4093. Each of the four states is individually accessible and can be used as either a normal NAND gate or an inverting Schmitt trigger by using the connections shown in Figure 2. All unused inputs of the package must be tied to the positive or negative supply rails, as appropriate.

Figures 3 to 5 show basic ways of using a 4093 gate as a Schmitt trigger. Each gate has a typical hysteresis voltage (difference between the upper and lower trigger threshold voltages) of 2 V when powered from a 10 volt supply. In

A	B	OUT
0	0	1
0	1	1
1	0	1
1	1	0
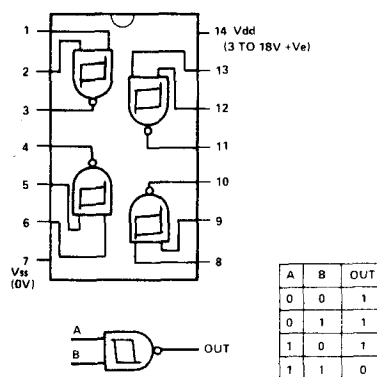
Figure 1 Functional diagram of the 4093 and operating truth table (for a single Schmitt gate).
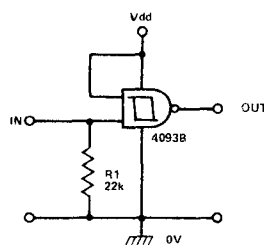
Figure 3. Simple Schmitt trigger with dc input.

Figure 3 the input signal is direct-coupled to the gate input. In the Figure 4 sine/square converter circuit the input signal is ac-coupled and the input pin is biased at half-supply via R1 and R2. In the improved sine/square converter circuit of Figure 5, the input pin bias can be adjusted to mid-way between the upper and lower threshold values, to give maximum sensitivity.
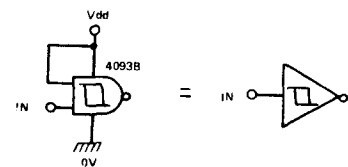
Figure 2. As the gates in a 4093 have a NAND function normally, they should be connected as above to operate as an inverting Schmitt trigger.
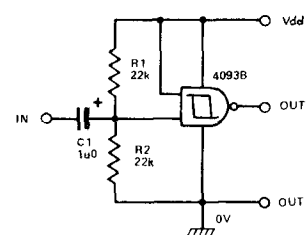
Figure 4. Simple connection for use as a sine/square converter.
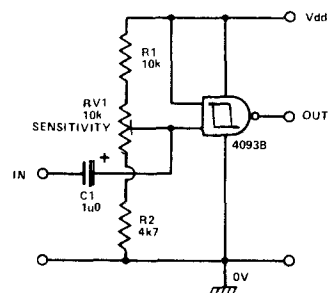
Figure 5. This circuit greatly improves the sensitivity when using the 4093 as a sine/square converter.
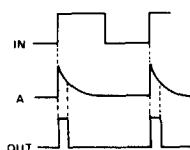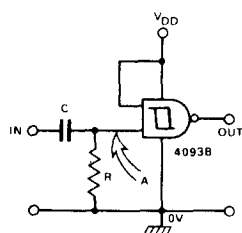
Figure 6. To trigger on the rising edge of a pulse, a CR network is added to the input.
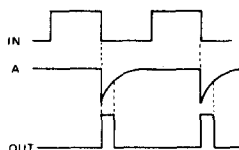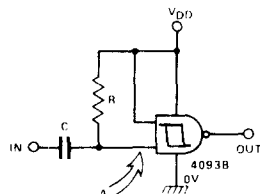
Figure 7. A re-arrangement of Figure 6 to produce a falling-edge trigger.
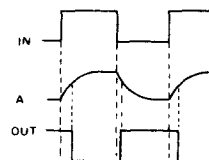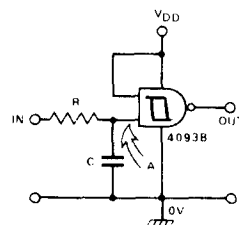
Figure 8. How to obtain a delayed and inverted version of an input pulse.
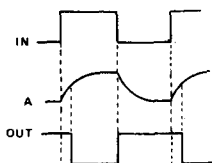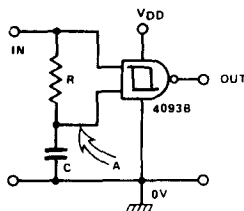
Figure 9. How to shorten a pulse by delaying the leading edge. Output is the inverse of the input.
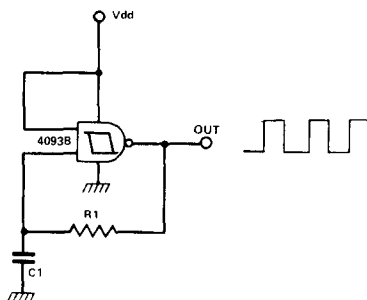
Figure 10. A basic astable multivibrator. The frequency of oscillation depends on the values of R1 and C1 and the hysteresis.
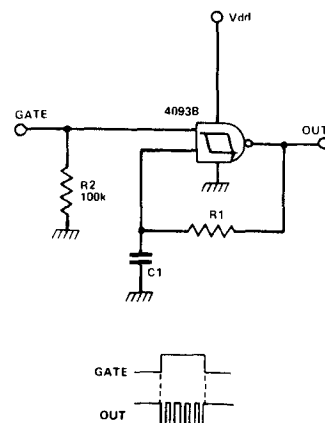
Figure 11. A variation on the circuit in Figure 10 allowing gated operation of the multivibrator.

Edge Detection

Figures 6 to 9 show a variety of ways of using the 4093 to detect or delay the edges of input pulse waveforms. The figure 6 circuit gives an output pulse on the arrival of the rising or 'leading' edge of an input pulse. The duration of the output pulse is determined by the CR values.

The Figure 7 circuit produces an output pulse on the arrival of the falling or 'trailing' edge of an input pulse. The Figure 8 circuit delays the entire input pulse by a period determined by the CR values. The circuit in Figure 9 delays the leading edge only.

Clock Circuits

Figure 10 shows how a single 4093 gate can be used as an astable multivibrator or 'clock' generator. This circuit gives excellent performance with very clean output edges that are unaffected by supply line ripple and other nasties. The operating frequency is determined by the CR values and can be varied from a few cycles per minute to 1 MHz or so. The circuit action is such that C1 alternately charges and discharges via R1. Capacitor C1 can be a polarized component.

Figure 11 shows how the basic astable can be gated on and off via an external

signal. Note that the circuit is gated ON by a high input, but gives a high output when it is in the OFF state.

The basic astable circuit of Figure 10 produces an inherently symmetrical output waveform. The circuit can be made to produce a non-symmetrical output by providing the timing capacitor with alternate charge and discharge paths, as shown in the circuits of Figure 12. This circuit produces fixed mark-space ratio output.

Figure 13 shows a special-purpose voltage-controlled astable which operates only when V_{in} rises above the upper Schmitt threshold: the operating
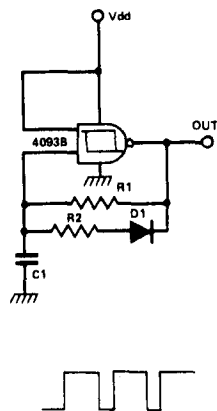
.Figure 12. **How to obtain a non-symmetrical mark-to-space ratio for the astable multivibrator.**

Figure 13. **The frequency of the astable may be varied by varying Vin shown here. However, Vin must be higher than the upper Schmitt threshold to start with. As you increase Vin, the frequency will increase.**

Figure 14. **"Debouncing" a press button switch. Such circuits are widely used in logic applications.**
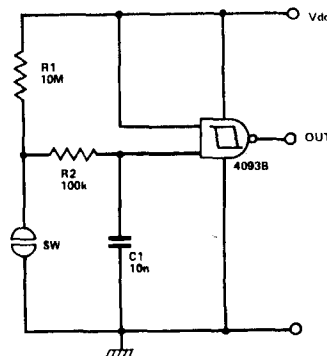
Figure 15. **A touch switch that provides a *high* output when operated. If you swap R1 and SW the output goes *low* when you touch SW.**
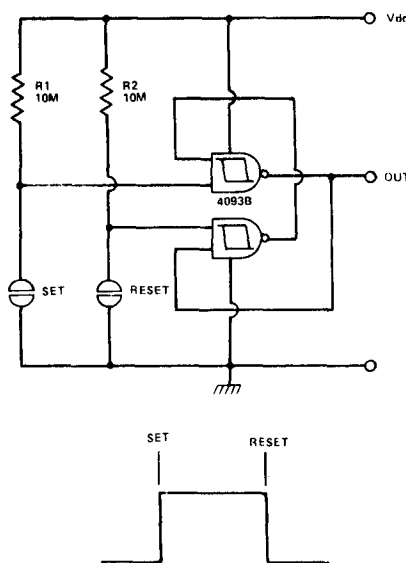
Figure 16. **A 'latching' or 'bistable' touch switch using two gates from the 4093.**

frequency then rises as V_{in} is further increased.

Miscellaneous

Figure 14 is the circuit of a 'noiseless' pushbutton switch, which produces a clean output pulse each time PB1 is operated. C1 charges up rapidly when PB1 is closed but discharges slowly with a period that is long (relative to normal noise spikes) via R1 when PB1 is released.

The output of the circuit in Figure 15 goes high when the input contacts are touched. A latching or 'bistable' touch-activated switch is shown in Figure 16.●

Lab Notes

Voltage regulators
— circuits and techniques

Here's a 'cookbooklet' of voltage regulator circuits and techniques covering everything from simple zeners to three-terminal regulators from UK correspondent **Ray Marston**, with additional material from staff.

ONE OF THE most common and mundane tasks facing the electronics enthusiast is that of designing voltage regulator circuits, ie: circuits that produce a stable and well defined dc output voltage over a wide range of load current variations. These circuits may vary from simple zener networks, designed to provide load currents up to only a few milliamps, to fixed-voltage high current units for powering logic boards, etc, or to variable-voltage high current units designed to act as general purpose pieces of test gear.
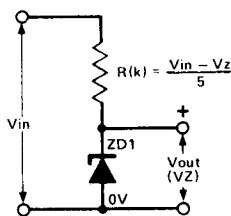
Figure 1. The simplest regulator. The zener regulator is often used to provide a fixed voltage 'reference' and supplies little or no load current. In such case a 400 mW zener of appropriate voltage is used and biased at about 5 mA. The value of R, in kilo ohms, can be calculated from the formula.

Zener based circuits

A zener diode can be used to produce a fixed reference voltage as shown in Figure 1. Often, the supply voltage (V_{in}) may be subject to fairly wide variations, causing the zener current to vary over a similarly large range. As long as V_{in} is always more than a few volts greater than the zener voltage and pro-

vided that the zener power rating is not exceeded this variation has only a moderate influence on the output voltage of the zener, which typically has an effective output impedance of a few tens of ohms.
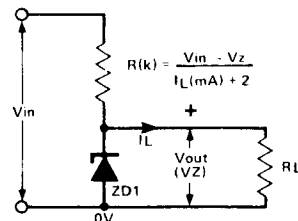
Figure 2. The basic zener regulator can be designed to supply a load current of a few milliamps to hundreds of milliamps depending on the zener power rating. Knowing the required load current, the value of R can be calculated from the formula.

A zener can be used as a very simple voltage regulator, providing load currents up to a few mA, by determining the value of 'R' as shown in Figure 2.
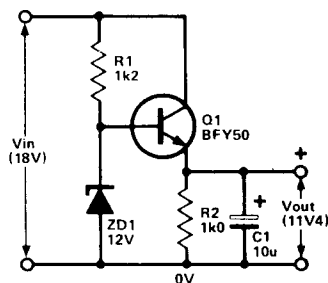
Figure 3. The basic 'series pass' fixed voltage regulator using a zener reference. This circuit can deliver up to about 100 mA at 11.4 volts. Q1 is basically an emitter follower, the base being 'clamped' at 12 V by the zener. C1 reduces the circuit's ac impedance.

In most practical voltage regulator applications the zener is simply used to apply a 'reference' voltage to a high-gain non-inverting buffer amplifier, which then supplies the required output power. The simplest example of this type of circuit is shown in the series-pass regulator circuit of Figure 3.

Op-amp regulators

One way of improving the regulation of the Figure 3 circuit would be to use a Darlington or super-alpha pair of transistors in place of Q1. An even better solution is to use the op-amp plus transistor buffer stage shown in Figure 4. Choice of circuit would perhaps depend on component availability and the application. ▶
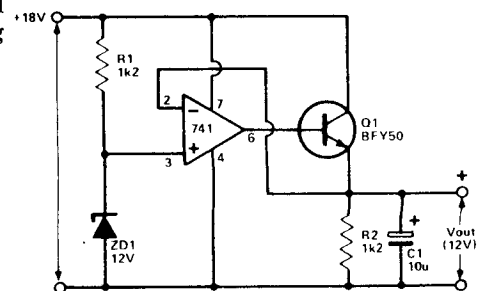
Figure 4. A considerable improvement in regulation is achieved by employing a high gain op-amp to drive the series pass transistor plus negative feedback from the dc output. This circuit will deliver up to about 100 mA and the output voltage stays within a few millivolts of the zener reference for output current variation from zero to maximum.

125

Lab Notes

The op-amp based regulator of Figure 4 supplies 12V at currents up to 100 mA and gives excellent regulation, despite its simplicity. The op-amp and transistor (Q1) are interconnected as a unity voltage gain non-inverting dc amplifier that has an overall current gain of about one million! The output voltage will be within a few millivolts of the zener reference value and the output impedance is less than one milliohm. The safe output current driving capacity of the circuit is limited to about 100 mA by the power rating of Q1, which should have a small 'flag' or push-on heatsink attached.
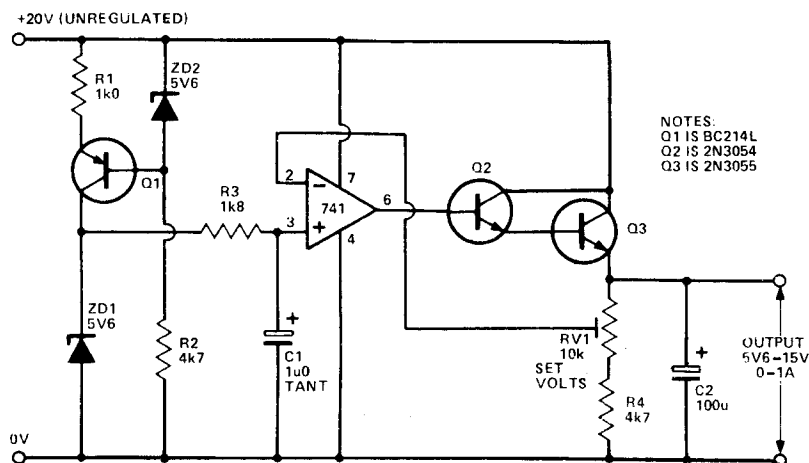
Figure 5. **This circuit is a little more sophisticated and performs better than the Figure 4 circuit. This circuit will deliver up to 1 A output current, output voltage being variable between 5.6 V and 15 V. Thermal drift is near zero.**

Figure 6. **This variation on the Figure 5 circuit permits the output voltage to be varied between 0 V and 20 V. The 3140 op-amp has a fixed voltage gain of about four, fixed by R4/R5.**

This circuit can be modified to provide greater current output by replacing Q1 with a power Darlington transistor such as the BDV65B. A suitable heatsink should be used.

The performance of the basic op-amp circuit of Figure 4 can be improved in a variety of ways, some of which are shown in Figure 5. The first improvement that can be made is to make the zener reference voltage (ZD1) independent of supply voltage variations by powering ZD1 from the output of constant-current generator Q1-R1-ZD2-R2. Next, the zener noise can be eliminated by feeding the reference voltage to the op-amp via low-pass filter R3-C1. Thermal drift effects can be

eliminated by making ZD1 a 5V6 type, with near-zero (actually about -0.2 mV/°C) temperature coefficient; the op-amp output voltage can then be set to the desired value (greater than 5V6) by using feedback components RV1-R4 to set the op-amp's voltage gain at some appropriate value.

Finally, the load current capacity of the circuit can be set to a fairly high value by using a Darlington-connected power transistor in the series pass output stage. Note that Q2 and Q3 could be replaced by a BDV65B Darlington device. ▶

Figure 7. **A further refinement is the addition of current limiting and overload protection. Q4 senses the voltage across R8 which is in series with the output. When the current exceeds 1 A, Q4** will be biased on and rob current from the base of Q2, preventing further increase in output current. The op-amp is protected against excessive supply rail voltage via R4 and ZD3.

Three terminal regulators

Three terminal regulators are remarkably easy to use, as shown in the basic circuits of Figures 8 to 10, which show the connections for making positive, negative and dual regulator circuits respectively; the ICs shown in these examples are 12 V units with current ratings of one amp. Note that a 270n or greater value disc ceramic capacitor should be connected close to the input terminal of the IC and a 10μ or greater electrolytic to the output.
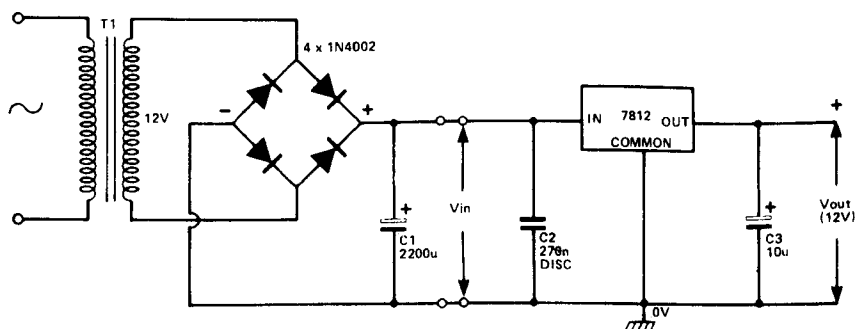
Figure 8. Circuit for the common-or-garden three-terminal fixed voltage positive regulator. Note that C2 and C3 should be mounted as close to the regulator's terminals as is physically possible.
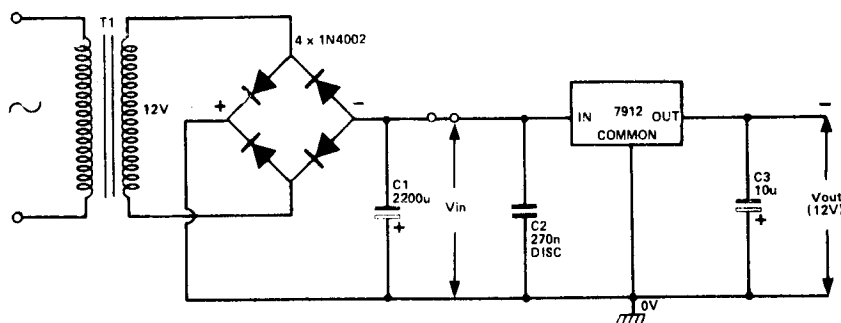
Figure 9. The negative voltage rail counterpart of Figure 8. Don't forget heatsinks.

These capacitors should be connected as close to the device's terminals as is physically possible. They can be quite good RF oscillators if this is not done correctly!

The output voltage of a three terminal regulator is referenced to the 'common' terminal of the IC, which is normally grounded. Most regulator ICs draw quiescent currents of a few mA, which flow to ground via the common terminal. The regulator output voltage can thus be 'jacked up' from the designed value by simply biasing the 'common' terminal with a suitable voltage.

There are two simple ways to do this. The first is shown in Figure 11. The required extra voltage is obtained by passing the IC's quiescent current through a trimpot (RV1). The trimpot is then set to provide the required output voltage. Quiescent current is typically 7 mA to 11 mA. This technique is adequate in most applications, although

the output voltage obviously shifts slightly with changes in the IC's quiescent current and thus regulation is degraded a little.

The output voltage of a three terminal regulator can be increased by a fixed amount by connecting a suitable value zener in series with the common terminal as shown in Figure 12. Output voltage is then the regulator voltage plus the zener voltage, in this case, 17.6 volts. Note that the quiescent current from the common terminal of the IC may not be sufficient to drive the zener and extra bias current is best provided. Here, R1 serves that purpose.
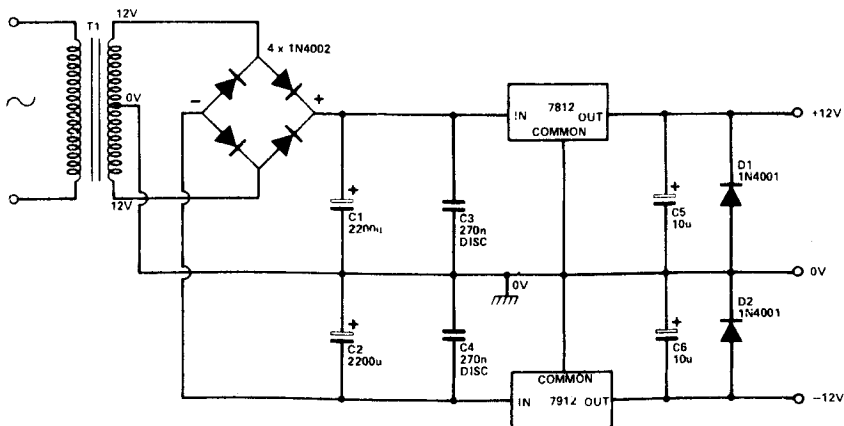
Figure 10. Circuit of a dual positive/negative power supply employing 1 A three-terminal regulators.
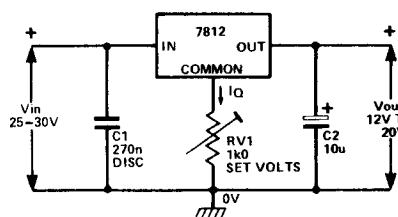
Figure 11. A very simple method of varying the output voltage of a three-terminal regulator. Regulation is slightly affected.
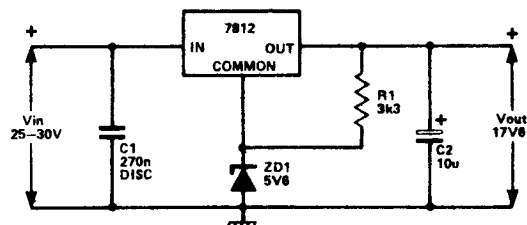
Figure 12. The output voltage can be 'jacked up' by inserting an appropriate zener in series with the regulator's common lead.

127
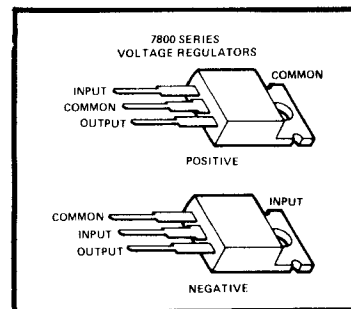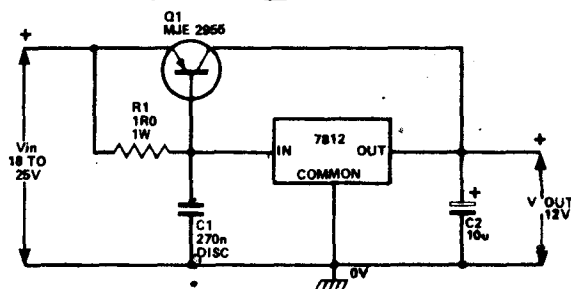
Figure 13. The output current capability of a three-terminal regulator can be boosted by using an external series pass transistor as illustrated. This circuit can deliver up to 5 A.
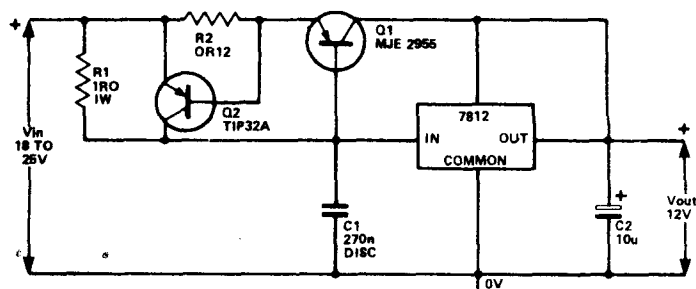

Figure 14. Adding current overload protection to the Figure 13 circuit. It's a wise idea, even though three-terminal regulators generally incorporate thermal shutdown protection.

More current

The output current capacity of a three-terminal regulator can readily be boosted via an external transistor as shown in Figure 13. This circuit, although it looks familiar, is a little unusual. At low load currents, less than 0.6 volts is developed across R1 and Q1 will not turn on. Thus, all the load current will be supplied by the 7812 regulator. When the load current is greater than about 600 mA, the voltage developed across R1 will be sufficient to turn Q1 on and Q1 will commence delivering current to the load.

This circuit can supply up to 5 A. Note that Q1 will need to be mounted on a substantial heatsink, capable of dissipating 60W.

Protection

Overcurrent or overvoltage protection for three terminal regulators can be provided relatively easily. Overcurrent protection for the Figure 13 circuit is illustrated in Figure 14. All it needs is a transistor and a resistor. When the current through R2 (a 0.12 ohm resistor) exceeds 5 A, the voltage across it will turn on Q2, robbing base current from Q1, thus current limiting the output. The 7812 will go into its switch-on/switch-off protection mode.

To protect a three terminal regulator from damage that may result from excessive supply line voltage, a 'pre-regulator' circuit is added. This is illustrated in Figure 15. You will recognise components Q1, R1, C1 and ZD1 as a simple zener-based series pass regulator (as in Figure 3).

Three terminal regulators can withstand no more than 40 V maximum between the input and common terminals, so this circuit limits the voltage applied to about 27 V. Note that Q1 should be mounted on a heatsink capable of dissipating 25 W.

The LED as a regulator

A low noise regulator can be made using a LED as the reference element. David Tilbrook used this technique in the Series 4000 moving-coil preamp (ETI-473, Oct '79).

The usual reference element for discrete voltage regulators is a zener. As these devices operate in the reverse-biased mode, they are inherently noisy and will put noise on the regulated supply rail. This is likely to degrade the performance of low noise, low level circuits supplied by the regulator.

The circuit of the ±6 V regulated supply for the ETI-473 moving-coil pre-amp is reproduced here. A red LED operated in the forward-biased mode drops a constant 1.65 V and generates very little noise. The reference LEDs in

the circuit here are LED1 and LED2. Series regulators Q13 and Q14 regulate the incoming ±12 V. The potential dividers R21/R23 and R22/R24 divide the voltage present at the output of the regulators and drive transistors Q15 and Q16, and the LEDs. The base-emitter junction in series with each LED will drop 0.6 V; to this is added the LED forward drop of 1.65 V. Thus, whenever the voltage present at the junction of the voltage divider resistors tries to increase above 2.3 V, Q15 and Q16 will increasingly conduct, decreasing drive to the bases of Q13 and Q14 respectively.

Noise on the regulated supply rails is further reduced by the C-R networks, C12/R19 and C13/R20. ●

Figure 15. Three-terminal regulators generally have a maximum input voltage rating of 40 V. Where the input may vary widely, at times exceeding the device's limit, this circuit provides protection by using a 'pre-regulator'.
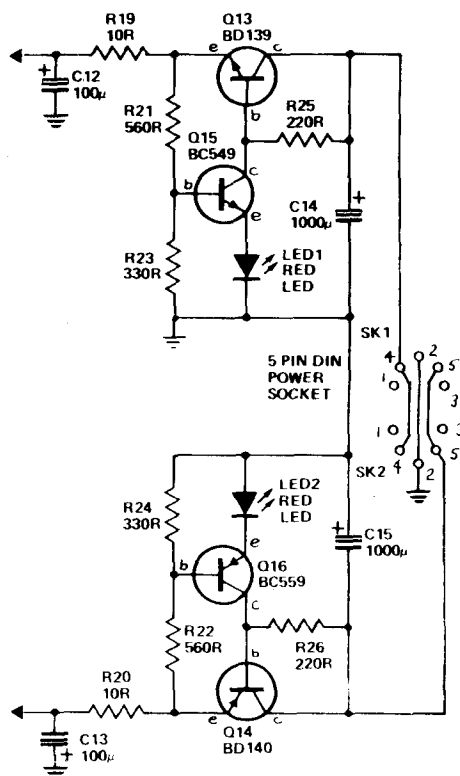
Figure 16. A low noise regulator is necessary in some applications, in which case zeners cannot be used as they generate noise. LEDs can be used as 'quiet' voltage reference sources, as illustrated in the +/- 6 V regulators from David Tilbrook's Series 4000 Moving Coil Preamp project.

Lab Notes

Designing with diodes
— tricks of the trade

The humble diode can extricate you from some difficult circuit problems, be it in logic or linear applications.

Ray Marston

THERE ARE numerous occasions when one needs a basic gate circuit or two and is faced with the possibility of having to wastefully commit an entire IC to this simple function. Alternatively, it may be the case that the inputs to a gate come from such widely separated points of a circuit that the use of an IC in a particular application will result in an excessively complicated pc board layout. In both of these instances, a simple diode gate may offer an ideal solution to the problem.

Figure 1 shows the practical circuit of a three input diode OR gate. The circuit is simple, reasonably fast, very cost-effective and can readily be expanded to accept any number of inputs by merely adding one more diode to the circuit for each new input.

The diode OR gate can be converted to a NOR type by either feeding its output through an NPN transistor inverting stage, as shown in Figure 2a, or by feeding its output through any type of IC inverting stage that happens to be 'spare' in the circuit you are playing with, as shown in Figure 2b.

Figure 3 shows the connections for making a three input diode AND gate. The circuit can again be expanded to accept virtually any number of inputs by simply adding an appropriate number of diodes.

The AND gate can be converted to a NAND type by feeding its output through a pnp transistor or an IC inverting stage, as shown in Figures 4a and 4b respectively.

Figure 2. The diode OR gate can be converted to a NOR type by feeding its output through a transistor, as in (a) above, or through an IC inverting stage, as in (b) below.

Figure 1. The diode OR gate is simple, but efficient. It can be expanded to accept any number of inputs by adding extra diodes.
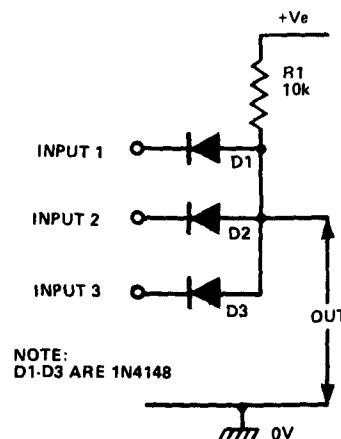
Figure 3. The circuit of a three-input AND gate. The number of inputs can be increased by adding extra diodes.
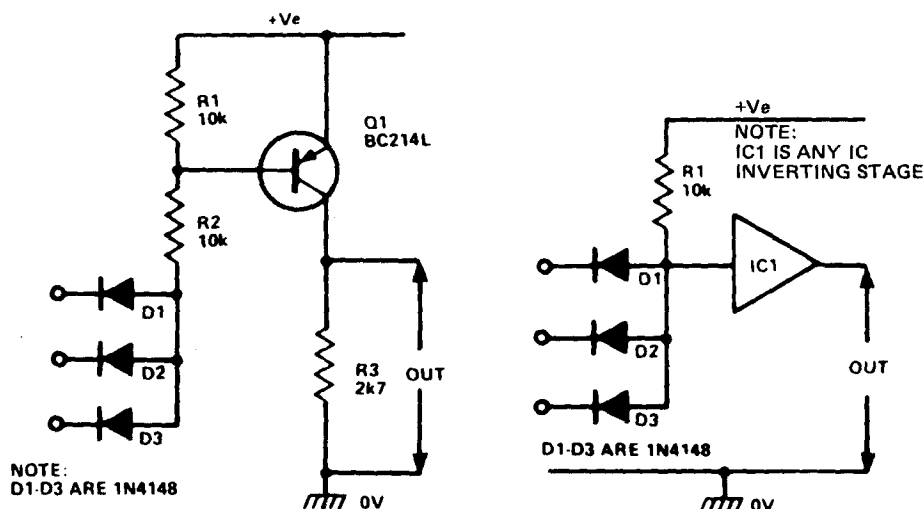
Figure 4. The diode OR gate can be converted to a NAND type by feeding its output through an inverting stage using either a transistor, as in (a) at left, or an IC as in (b) at right.

Diode volt drops

We've mentioned above that the output of the 'analogue' diode gate may be 'within a diode volt drop' of the input signal. The magnitude of this 'volt drop' depends on the type of diode that is in use, on the magnitude of the diode forward current and on the temperature of the diode junction. All silicon diodes have a negative temperature coefficient of about -2 mV/°C.

Figures 9 and 10 show typical volt-drop curves for the popular 1N4148 and 1N4001 silicon diodes at 25°C. The graph of Figure 9 spans the current range 100 uA to 1 mA and the graph of Figure 10 spans the range 1 mA to 50 mA.

Note that the 1N4148 volt drop typically ranges from 519 mV at 100 uA to 874 mV at 50 mA, compared to the 1N4001's range of 441 mV at 100 uA to 744 mV at 50 mA.
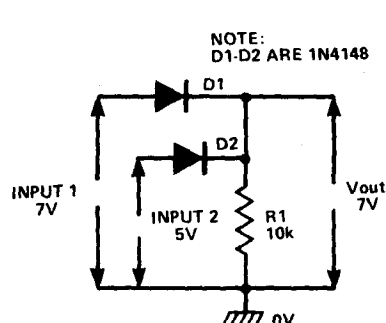
Figure 5. When a diode OR gate is used in the linear mode, Vout equals the greater of the two inputs.
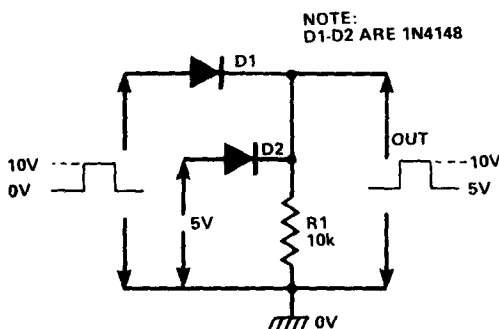
Figure 6. The effect of feeding a pulse to one input and a dc voltage to the other input of a two-input diode OR gate.
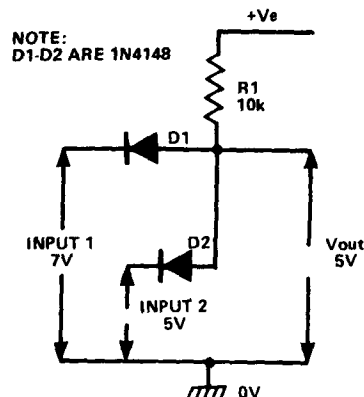
Figure 7. When a diode AND gate is used in the linear mode, Vout equals the less of the two inputs.

Linear operation

Diode AND and OR gates can be very useful when one or more of their inputs are operated in the linear mode. Figures 5 and 6 show two useful ways of using the two-input diode OR gate in linear applications

In the case of the Figure 5 circuit, in which analogue voltages are applied to both of the input terminals, the output of the circuit (ignoring a diode volt drop of about 600 mV) equals the greater of the two input voltages.

Figure 6 shows what happens when a pulse signal is fed to one input of the OR gate and an analogue voltage is fed

to the other. The output signal comprises a pulse with a peak amplitude equal to that of the input pulse and with a 'zero' value equal to the analogue input voltage.

Figures 7 and 8 show similar circuits based on the two-input diode AND gate. In the Figure 7 circuit, where analogue voltages are fed to both inputs, the output (ignoring a diode volt drop 'gain' of about 600 mV) equals the lesser of the two inputs.

In the case of the Figure 8 circuit, where a pulse is fed to one input and an analogue voltage to the other, the output pulse has a peak amplitude equal to that of the analogue input voltage.
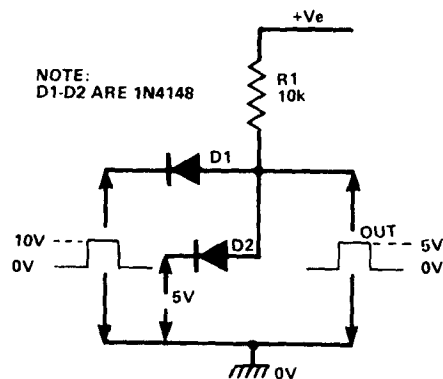
Figure 8. The effect of feeding a pulse to one input and a dc voltage to the other input of a two-input diode AND gate.
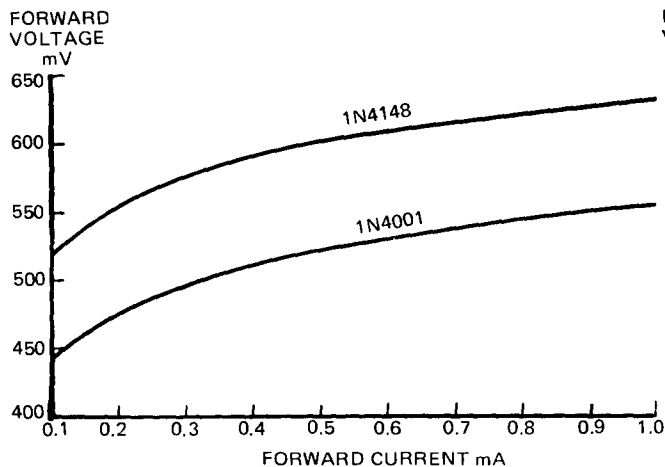
Figure 9. **Diode volt-drop curves for the 1N4001 and 1N4148 diodes over the 100 uA to 1 mA current range.**
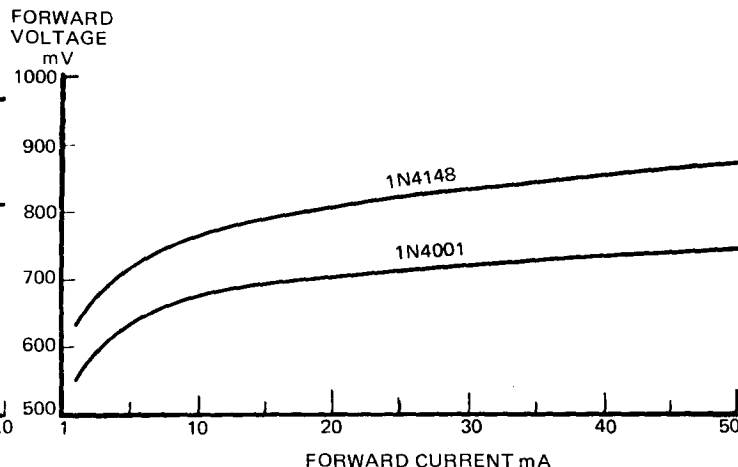
Figure 10. **Diode volt-drop curves for the 1N4001 and 1N4148 diodes over the 1 mA to 50 mA current range. Note that these two diode types are the most commonly used (highest volume sales of any single component).**

A point of particular note about the 1N4001 curve is that its volt drop of 714 mV at 25 mA increases by only a fraction over 4% (to 744 mV) when the forward current is doubled, to 50 mA. In other words, the diode has a voltage-to-current coefficient of about 0.04%/% in this current range. The diode can thus be used as a reasonably stable voltage reference at these current levels but has a negative temperature coefficient of about −0.3%/°C.

Constant current generator

Figure 11 shows how the 'voltage reference' characteristics of the 1N4001 can be put to good use in a simple constant-current generator circuit that can be used for re-charging Ni-Cad cells or for linearly charging large capacitors, etc. Here, two 1N4001s are wired in series

and operated at a current level of roughly 50 mA. Consequently, the voltage across Rx is equal to the volt drop of the two diodes minus the base-emitter volt of Q1 (about 700 mV), which gives an Rx voltage of about 700 mV. The emitter (and hence collector) current of Q1 is thus approximately 700/Rx in (mA).

To give an idea of the magnitude of things, an Rx value of 1R2 (1.2 ohms) gives an output current of about 600 mA, 3R9 (3.9 ohms) gives about 200 mA, and 6R8 (6.8 ohms) gives about 100 mA. All in all, a simple but very useful circuit.

Diode protection circuits

To wrap up these Lab. Notes, let's take a quick look at some diode protection circuits. By 'protection' we mean circuits that are designed to insure devices against irreversible damage and

also circuits that are designed to prevent simple malfunctioning. Figures 12 to 15 show four circuits in this latter category.

In the case of Figure 12, we have a basic time constant circuit in which a rising voltage, with a time constant of about 100 seconds, is developed across C1 each time SW1 is closed. This voltage may be used to activate some additional circuitry. The problem is that, once C1 has charged up it has no means of rapidly discharging again (resetting) once SW1 is opened. If there is a load in parallel with the C-R network, as shown dotted in the diagram, C1 will of course discharge via R1 and the load but then has a very long time constant (greater than 100 seconds).

An easy way round this problem is to connect a discharge diode in parallel with R1 as shown in Figures 12 and 13.▶
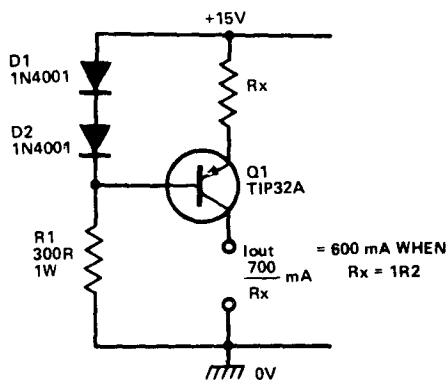
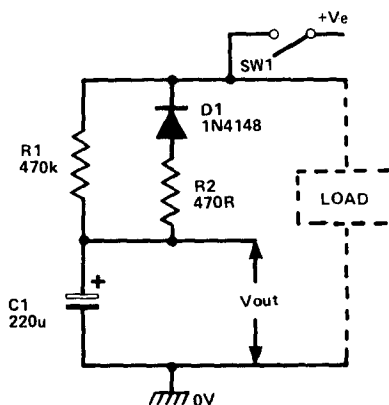Figure 11. **A simple and very useful current generator.**

Figure 12. **An example of the use of a diode to rapidly discharge a timing capacitor when the power supply connection is broken.**
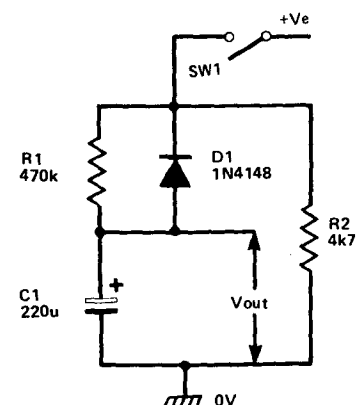

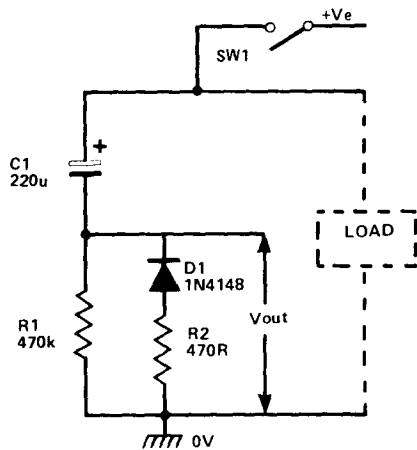Figure 13. **A modification to the Figure 12 circuit.**

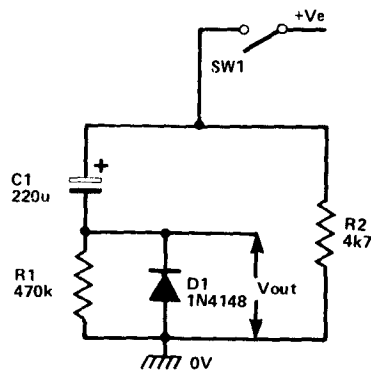Figure 14. A basic variant of the Figure 12 circuit.

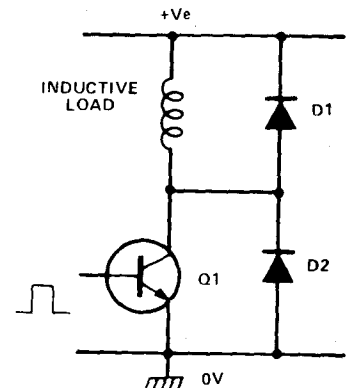Figure 15. A basic variant of the Figure 13 circuit.

Figure 16. An example of the use of diodes to protect a pulse-driven common emitter amplifier with an inductive collector load.

If there is a low-impedance load in parallel with the C-R network a current-limiting resistor must be wired in series with the discharge diode, as shown in Figure 12. If there is no load in parallel with the C-R network then an artificial load must be provided to complete the discharge path, as shown in Figure 13.

Figures 14 and 15 show two basic variations of the above circuits in which the C and R networks are configured to give a falling output voltage across R1.

Finally, Figures 16 and 17 show ways of using diodes to protect two types of transistor circuit from destructive damage. Figure 16 shows how to protect a pulse-driven common-emitter ampli-
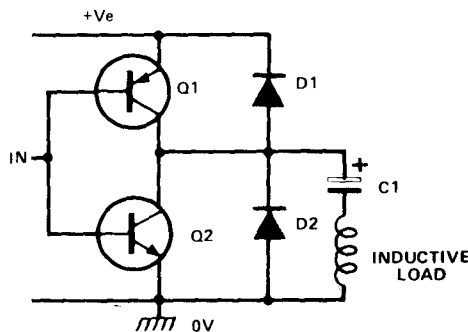
Figure 17. An example of the use of diodes to protect the complementary emitter follower output stage of a power amplifier that is used to drive an inductive load.

fier that has a highly inductive collector load, such as a transformer or a relay coil. Very high back EMFs can be generated by inductive loads and can easily

destroy transistor junctions. In the diagram D1 prevents the collector of Q1 from being driven above the positive supply rail value by these back EMFs and D2 prevents it from being driven negative.

Figure 17 shows how a similar type of protection can be given to the complementary emitter follower output stages of a power amplifier used to drive highly inductive loads. This circuit can give good protection to hi-fi amplifiers in which the speakers may be inadvertently plugged in at a moment when the amplifier is being hard driven. The protection diodes must have a current rating that is compatible with the inductive (speaker) load. ●

Lab Notes

A little light on LEDs

Ray Marston has provided some interesting notes on LEDs and how to use them.

Basic characteristics

SOMETHING that we all know about the LED is that it glows a pretty colour if we shove a bit of current through it. LEDs are presently available in four colours, red, orange, yellow and green. Blue LEDs will also be available in the near future. A voltage is developed across the LED when it is passing a forward current. Figure 1 shows typical forward voltage of different coloured standard LEDs at forward currents of 20 mA.

COLOUR	RED	ORANGE	YELLOW	GREEN
VF (TYPICAL)	1.8V	2.0V	2.1V	2.2V

Figure 1. Typical forward voltage characteristics of standard LEDs with forward current set at 20 mA.

When you use an LED, you have to wire some form of current-limiting device in series with it. Usually, a resistor can be used for current limiting. Figure 2 shows how to work out the value of resistance to give a particular current from a specific supply voltage: in practice, 'R' can be connected in either the anode or cathode side of the LED. The higher the operating current, the brighter the LED will glow. Most LEDs will operate safely up to absolute maximum currents of 30 to 40 mA.

You can use an LED as an indicator in an ac circuit by wiring a diode in inverse parallel with it, as shown in
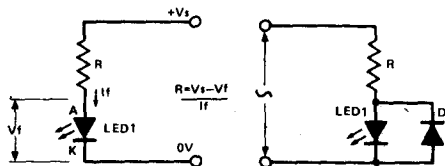
Figure 2. (Left). Finding the required series resistance from a known supply voltage.
Figure 3. (Right). Using a LED as an indicator in an ac circuit.

Figure 3, to prevent the LED being reverse-baised. For a given brightness, the value of 'R' should be halved relative to that of a dc circuit.

If an LED is reverse biased, it will avalanche or 'zener' at a fairly low voltage, as shown in Figure 4. Most LEDs have maximum reverse-voltage ratings in the range three to five volts. These low ratings present a trap for the unwary user, so take heed.
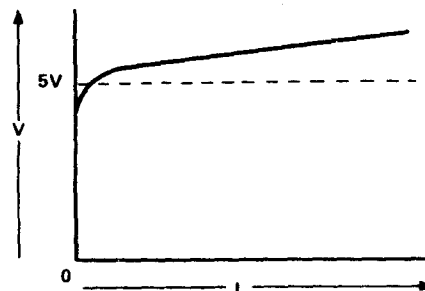
Figure 4. A reverse-biased LED will have characteristics similar to a zener with a 'knee' around five volts.

Pitfalls

The first practical problem that you'll encounter when using an LED is that of identifying its polarity. Most LEDs have their cathode identified by a notch or flat on the package, or by a short lead. This practice is not universal, however, so the only sure way to identify an LED is to test it in the basic circuit of Figure 2: try the LED both ways round. When it glows, the cathode is the most negative of the two terminals. It is always good practice to test an LED before soldering it into circuit.

The second pitfall concerns the use of those 'cheapo' LEDs that come in Bargain Packs. These are usually advertised as 'second grade' or 'out of spec' devices, but just how out-of-spec they are can sometimes be quite mind blowing. You'll often find that half of the devices in a pack have forward voltages in the range five to eight volts, which makes them virtually useless in many applications.

If you ever need to drive a number of LEDs from a single currently available 'dot' or 'bar' LED-display driver ICs, always check its spec to see if it is sensitive to LED characteristics. The Siemens UAA170 15-LED 'dot' driver, for example, will only function correctly if all LED forward voltages are matched to within 0.5 volts, and can thus be used with first grade LEDs only. ▶
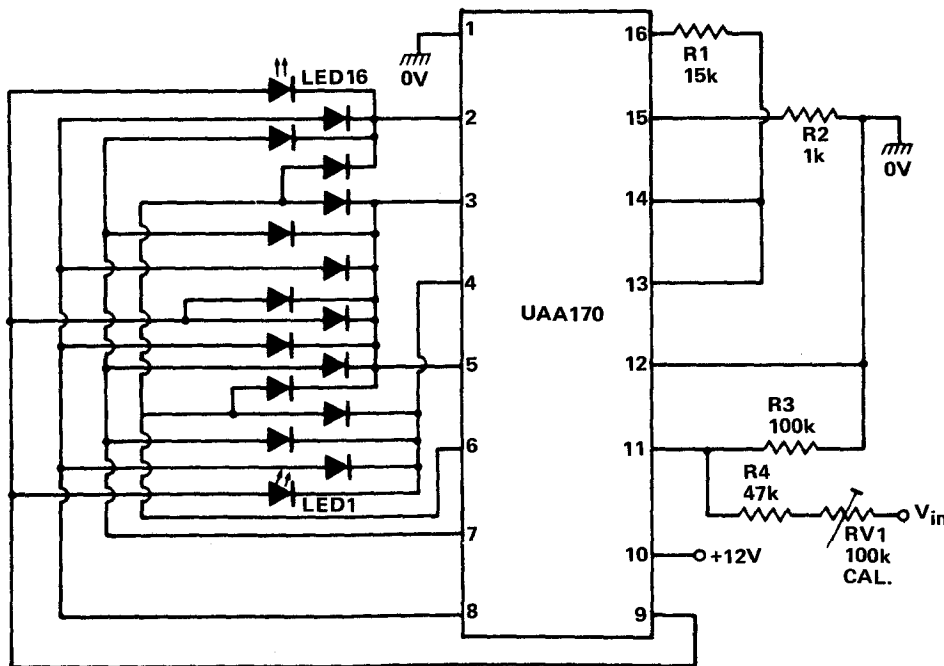
Figure 5 shows the circuit of a 0 to 10 volt 16-LED voltmeter using this IC.

Driven to it

If you ever need to drive a number of LEDs from a single source, take notice of the circuits in Figures 6 to 9. Figure 6 shows how a number of LEDs can be wired in series and driven via a single current-limiting resistor. Note that the supply voltage used here must be significantly greater than the sum of the individual LED forward voltages. This circuit thus draws minimal total current, but is limited in the number of LEDs that it can drive.

The Figure 7 circuit, on the other hand, can drive an unlimited number of LEDs, but is very wasteful of current. The total current drawn is equal to the sum of the individual LED currents.

Figure 8 combines the Figure 6 and 7 circuits to give the best of both worlds. The circuit can drive an unlimited number of LEDs, at maximum current economy.

Figure 9 illustrates one of those 'traps for the unwary', or 'what not to do' circuits. This circuit will not function correctly because inevitable differences in the forward voltage characteristics of the LEDs will usually cause one LED to 'hog' most of the available current, leaving little or none for the remaining two.

Chasers

The highly popular CD4017B decade counter with 10 decoded outputs is widely used for driving LED displays in chaser or sequencer applications. A certain amount of confusion seems to exist, however, concerning the 'correct' method of connecting the LEDs to the decoded ouputs.

The decoded outputs of this CMOS device provide inherent current-limiting under short-circuit conditions. The manufacturers do not quote a maximum short-circuit current value, but practical experience indicates that currents of 10-10-15 mA are commonly available from the 'B' version of the 4017. A maximum device dissipation per output transistor figure of 100 mW is quoted on some data sheets, indicating that a volt drop
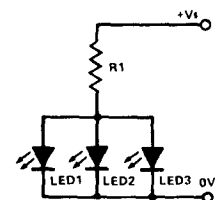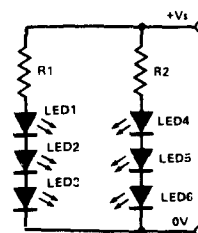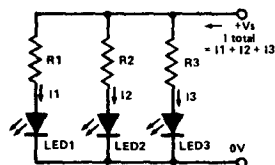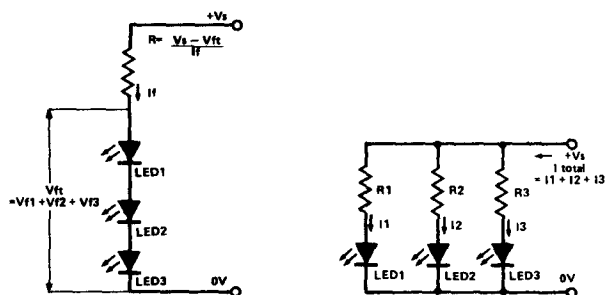

Figure 6. (Left). How to determine the series resistor value required for LEDs wired in series and driven from a single supply.
Figure 7. (Right). You can drive a whole host of LEDs from one supply rail — provided you can source the current required.

Figure 8. (Left). This is a combination of the circuits in Fig's 6 & 7.

Figure 9. (Right). How NOT to do it — one LED hogs all the current.

up to about seven volts can safely be developed across a 4017 output stage under maximum-current conditions.

Thus the LED chaser circuit of Figure 10, this has each LED connected directly between an output and ground and can safely be used up to maximum supply values of 9 volts. At voltages greater than 9 volts, the circuit of Figure 11, which has a resistor wired in series with each LED, should be used. Note that the main purpose of these resistors is that of reducing the power dissipation of the 4017B.
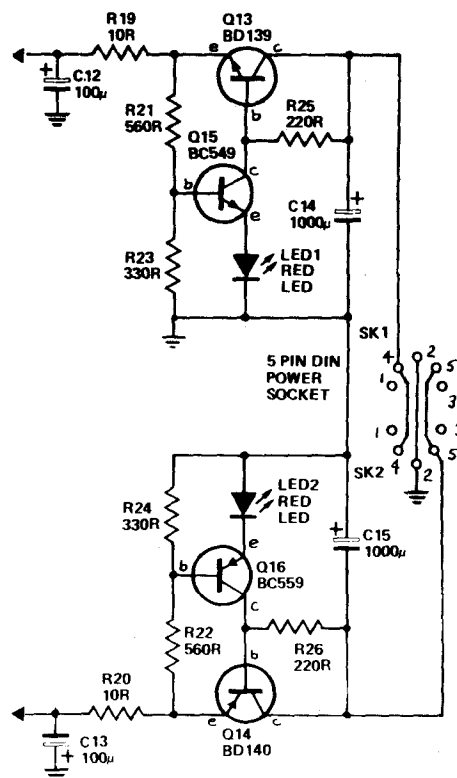
A variant that is sometimes used is shown in Figure 12a, and can be used with reasonable confidence at supply levels up to 12 volts maximum. Figure 12b shows a possible equivalent of this circuit when it is powered from a 15 volts supply, and illustrates the defect of the design. The action of the 4017 is such that when a given LED is ON, the anodes of all other LEDs are effectively grounded. R1 thus causes the OFF LEDs to be reverse biased. Because of the low reverse-voltage ratings of LEDs, it will often be found that one of the OFF LEDs will Zener at about five volts, giving the results shown in the diagram and possibly causing a destructive power overload in one of the 4017B output stages. Figure 12 thus represents a classic 'trap for the unwary' type of LED circuit. ●

The LED as a regulator

A LOW NOISE regulator can be made using an LED as the reference element. David Tilbrook used this technique in the Series 4000 moving-coil preamp (ETI-473, Sept '79).

The usual reference elment for discrete voltage regulators is a zener. As these devices operate in the reverse-biased mode, they are inherently noisy and will put noise on the regulated supply rail. This is likely to degrade the performance of low-noise, low-level circuits supplied by the regulator.

The circuit of the +/−6V regulated supply for the ETI-473 moving-coil preamp is reproduced here. A red LED operated in the forward-biased mode drops a constant 1.65 V and generates very little noise. The reference LEDs in the circuit here are LED1 and LED2. Series regulators Q13 and Q14 regulate the incomeing +/−12V. The potential dividers R21/R23 and R22/R24 divide the voltage present at the output of the regulators and drive transistors Q15 and Q16, and the LEDs. The base-emitter junction in series with each LED will drop 0.6 V; to this is added the LED forward drop of 1.65 V. Thus, whenever the voltage present at the junction of the voltage divider resistors tries to increase above 2.3 V,

Q15 and Q16 will increasingly conduct, decreasing drive to the bases of Q13 and Q14 respectively.

Noise on the regulated supply rails is further reduced by the C-R networks, C12/R19 and C13/R20.
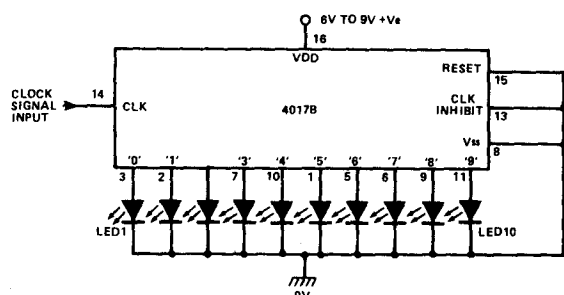

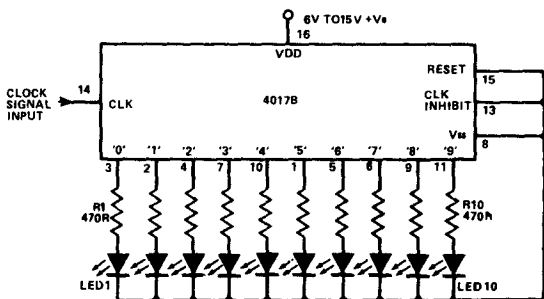Figure 10. A typical LED chaser circuit using a CMOS chip, the 4017B.

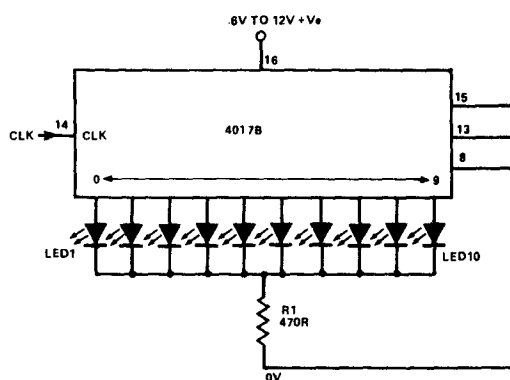Figure 11. For supplies over 9V, use this circuit instead of Figure 10.

Figure 12. The complete circuit at top may be used with supplies up to 12V but contains a trap for the unwary — at 15V, the 'off' LEDs will be biased into their zener region.

135

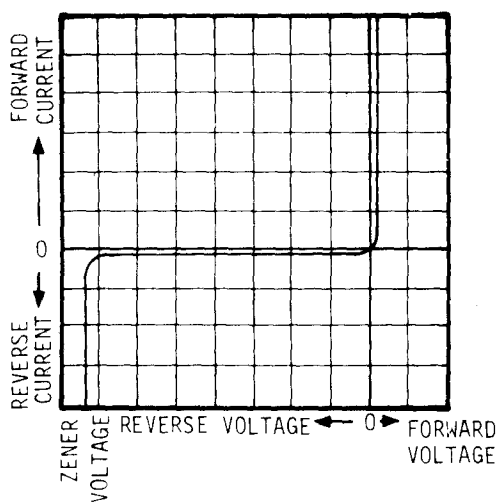# A PRACTICAL GUIDE TO ZENER DIODES

FIG. 1

In this two-part article we explain how to use this versatile circuit component

ZENER DIODES — named after Dr Carl Zener — are used primarily as voltage references. They are devices that maintain an almost constant voltage across them despite various changes in circuit conditions.

Externally, Zener diodes look much the same as conventional diodes and are manufactured in axial lead, and stud mounting forms; furthermore Zener diodes are capable of rectifying alternating current into pulsating direct current as are their conventional counterparts.

But unlike conventional diodes,

Zener diodes are deliberately intended to be used with the anode connected to a negative potential and the cathode connected to a positive potential.

When connected in this manner, Zener diodes have a very high resistance below a certain, critical, voltage (called, appropriately, the Zener voltage).

But if this critical voltage is exceeded, the dynamic resistance of Zeners drop to a very low level. And in this region, essentially constant voltages will be maintained across the Zeners, and these constant voltages will be maintained despite quite large changes in the applied currents.

This characteristic is illustrated graphically in Fig. 1 from which it may be seen that beyond the 'Zener voltage' the reverse voltage remains practically constant despite changes in reverse current.

Because of this characteristic, Zener diodes may be used to provide a constant voltage drop, or reference voltage, across their internal resistance.

Zener diodes are manufactured in a number of wattage ratings and with Zener voltages ranging from 2.7 Volts to 200 Volts. (In practice, Zener diodes with ratings exceeding 30 Volts are rarely used.)

THE BASIC CIRCUIT

A basic voltage regulator circuit, using one resistor and one (idealized) Zener diode, is shown in Fig. 2. In the example shown, the Zener diode is rated at 5.6 Volts and the applied
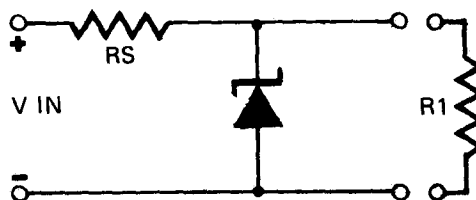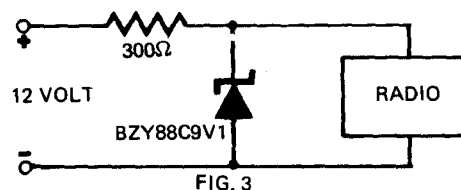

FIG. 2

FIG. 3

voltage V in is 8.0 Volts. With no load applied to the output terminals, 5.6 Volts will be dropped across the Zener diode and the remaining 2.4 Volts will be dropped across resistor R_s.

If the input voltage is now changed, say, from 8.0 Volts to 9.0 Volts, then the voltage drop across the Zener will still remain at its nominal 5.6 Volts but the voltage drop across R_s will increase, from 2.4 Volts, to 3.4 Volts. The voltage drop across the (ideal) Zener will remain constant.

In practice, the voltage will increase slightly due to the dynamic resistance of the Zener. (The method of calculating the change in Zener voltage is simply to multiply the dynamic resistance of the Zener diode by the change in Zener current.)

The resistor R1, in Fig. 2, represents an external load. If this load is connected across the circuit, then some of the current that was flowing through the Zener will now pass through the load. Providing the current in R_s is greater than the load current, some current will still pass through the Zener and a substantially constant voltage will be maintained across the Zener/load.

The series resistor R_s is selected so that the minimum current passing through the Zener· is not less than the level required for stable regulation.

This level begins just below the 'knee' of the reverse voltage/reverse current curve (as shown in Fig. 1). Curves such as this are available for all makes and types of Zener diodes.

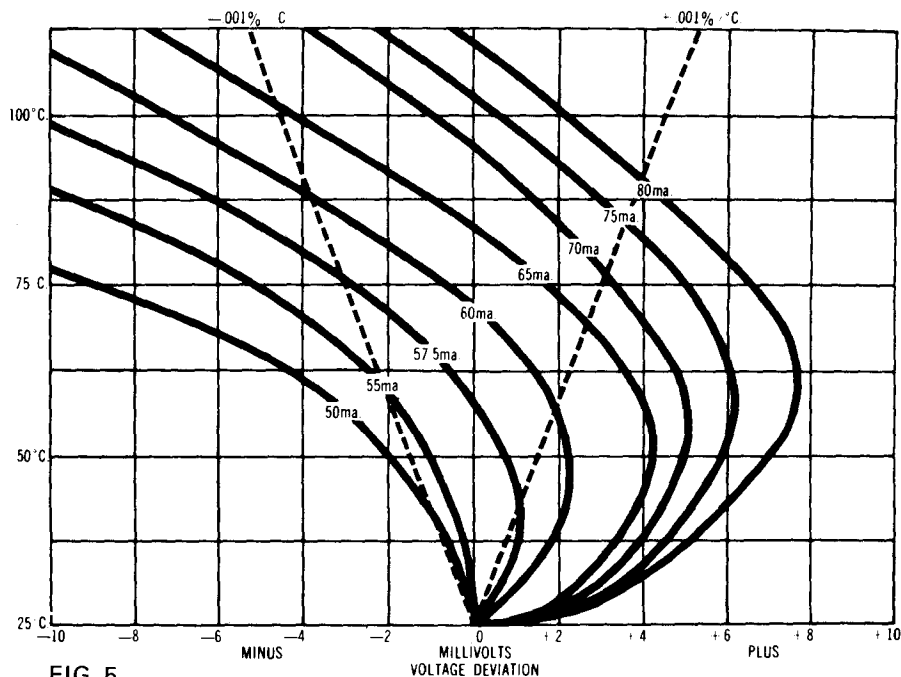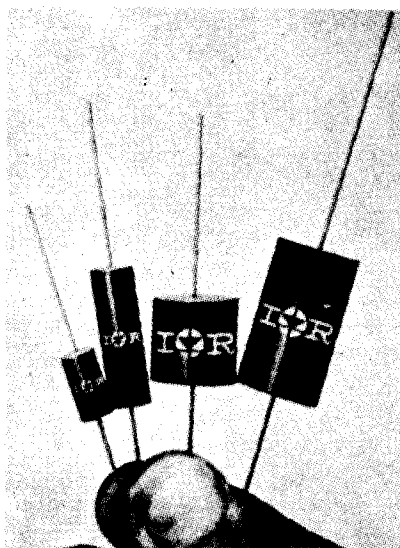It is also necessary to ensure that the value of R_s is such that current flow

FIG. 5

FIG. 4

through the Zener diode cannot exceed its specified power rating: which is the Zener voltage multiplied by the Zener current. In the circuit shown in Fig. 2, maximum current will flow through the Zener diode when the load is disconnected.

— The design procedure for this circuit is quite simple.

1. Specify the maximum and minimum load current (I_1), say 10 mA and 0 mA.
2. Specify the maximum supply voltage that is likely to occur (say, 12 Volts), but ensure that 'the minimum supply voltage will always be at least 1.5 Volts higher than the Zener voltage of the diode to be used.
3. In the circuit shown in Fig. 2, the required output voltage, and hence the Zener voltage, is 5.6 Volts, and the specified minimum Zener current is 100 microamps. Thus the maximum Zener current is 100 microamps plus 10 milliamps — which is 10.1 milliamps.
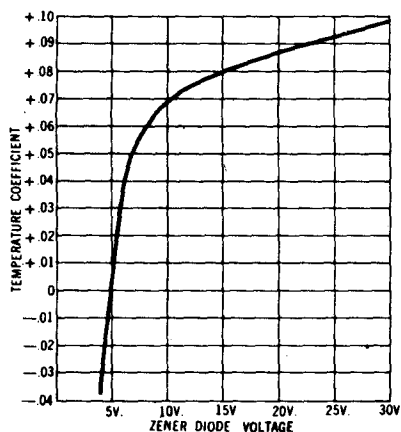
4. The series resistor R_s must conduct 10.1 mA at the lowest input supply voltage: and so allowing 1.5 Volts minimum voltage drop across R_s (i.e. input voltage minus Zener voltage):—

$$R_s = \frac{1.5}{10.1 \times 10^{-3}}$$
$$= 148.5 \text{ ohms}$$

5. The value of R_s is thus 148.5 ohms, and the nearest preferred value to this is 150 ohms.
6. At the maximum supply voltage (12 Volts), the voltage drop across R_s is $I_z R_s$, (I_z being Zener current):—

$$\text{thus } I_z = \frac{(12 - 5.6) \text{ mA}}{150}$$
$$= 42.6 \text{ mA}$$

7. This is the maximum current that will flow through the Zener at any time, i.e., maximum input voltage and zero external load. The power dissipated by the Zener under these conditions is:—

$$I_z V_z$$
$$= 5.6 \times 42.6$$
$$= 238 \text{ mW.}$$

8. Having calculated that the correct value for R_s is 150 ohms, and that the Zener diode must be capable of dissipating 238 mW all that remains is to choose the correct type of 5.6 Volt Zener. This is, in fact, well within the capabilities of most small Zener diodes which are generally rated at 400 mW.

A PORTABLE RADIO IN YOUR CAR

A practical example of the circuit described above is illustrated in Fig. 3. This shows how to run a transistor radio from your car's power supply.

The voltage required for the radio is 9 Volts — and the nearest Zener diode to this is 9.1 Volts.

The maximum current drawn by the radio is measured at — say — 10 mA, and, as the load does not vary greatly, this is also the minimum current. (As 10 mA is well above the 'knee' current of most small diodes, the addition of a minimum operating current is not required.)

The maximum supply voltage is around 14 Volts, (when the battery is fully charged).

Thus the series resistor R_s must conduct 10 mA at the lowest supply voltage (12 Volts) and its value is calculated by:—

$$R_s = \frac{12 - 9.1}{10 \times 10^{-3}}$$
$$= 290 \text{ ohms}$$

At the maximum supply voltage (14 Volts) the voltage drop across R_s is $I_z R_s$:—

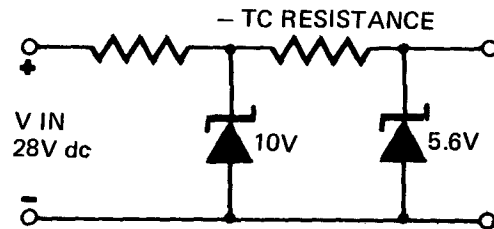$$I_z = \frac{(14 - 9.1)}{290}$$
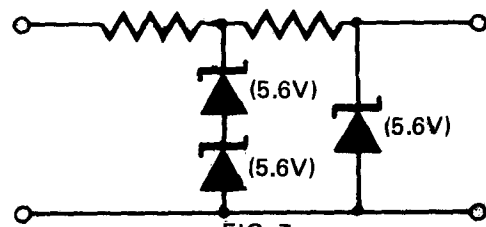$$= 16.8 \text{ mA}$$

FIG. 6

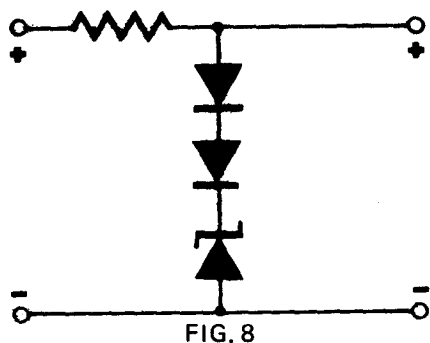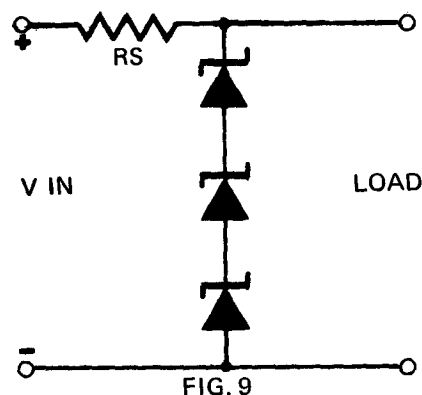FIG. 7

A PRACTICAL GUIDE TO ZENER DIODES

FIG. 8

FIG. 9

The maximum power dissipated in the Zener is:—

$$9.1 \times 16.8 \text{ mA}$$
$$= 152.8 \text{ mW}$$

This is well within the capabilities of a standard 400 mW Zener, such as a BZY 88 C9V1.

TEMPERATURE DRIFT

Although Zener diodes are relatively immune to changes in supply voltage and load, they are to varying extents affected by changes in ambient and operating temperature. Fig. 4 shows a typical Zener diode temperature coefficient curve, and, as can be seen, the coefficient, although approaching 0.1% per degree Centigrade at the higher voltages, passes through zero at about 5 Volts and then becomes negative for lower voltages, reaching -0.04%/C° at approximately 3.5 Volts.

The transition between a negative and a positive temperature coefficient is not well defined and a 5.6 Volt Zener may have either a positive or negative coefficient depending upon the current flow through it. But by careful control of the operating current it is possible to hold a standard 5.6 Volt Zener diode to a temperature coefficient of ± 0.001% over a temperature range of +25°C to +75°C (Fig. 5 refers).

The only difficulty is to provide a constant current source for the Zener diode. One method is to use a 10 Volt Zener diode to act as a regulator for the reference diode. (Fig. 6). The current limiting resistor between the two diodes should have a negative temperature coefficient to compensate for the positive (+0.07%/°C) temperature coefficient of the 10 Volt Zener.

A second method of stabilizing the current to the Zener diode is shown in Fig. 7. In this arrangement two 5.6 Volt Zeners, each having practically zero temperature coefficient, in turn stabilize the current through the voltage reference.

Another approach to compensation for temperature change is shown in Fig. 8 where the temperature coefficient of forward-biased silicon diodes is used to counteract the opposite temperature coefficient of the Zener diode. This method may be used to compensate for voltage drift over quite wide ranges of both temperature and voltage, Zener voltage stabilities of 0.01% are readily achieved.

As can be seen from Fig. 4, the temperature coefficient of Zener diodes increases considerably at the higher Zener voltages. An alternative method of temperature stabilizing to that shown in Fig. 8, is to use a number of 5.6 Volt Zeners in series. (Fig. 9).

TEMPERATURE SENSING

The apparent disadvantage of a Zener diode's temperature coefficient may be put to a useful purpose in the form of a temperature sensing device. Fig. 10 shows how a bridge consisting of two resistors and two similar Zener diodes may be constructed so as to indicate a temperature level when one of the diodes is held at a reference temperature and the other is subject to the conditions to be monitored. The average small 10 Volt Zener has a temperature coefficient of +0.07%/°C — this corresponds to 7 millivolts per °C change. The sensing element will, therefore, indicate an imbalance of 0.7 Volts when undergoing a 100°C temperature change.

NON-STANDARD VOLTAGES

Occasionally it is necessary to obtain a regulated voltage other than that obtainable from a single Zener, this
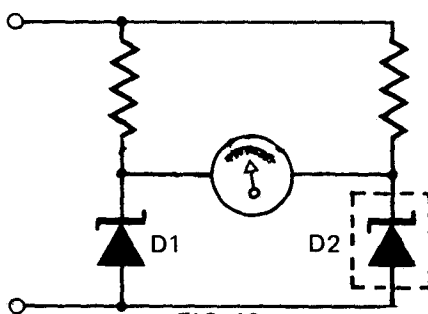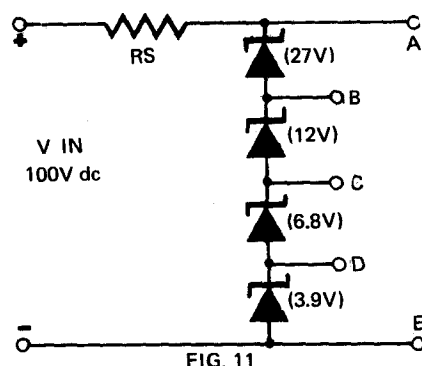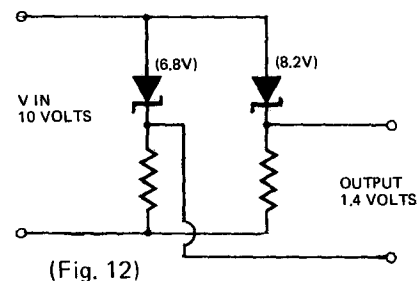
FIG. 10

FIG. 11

(Fig. 12)

can be achieved by connecting a number of Zener diodes in series (Fig. 9). The diodes need not have equal breakdown voltages since the arrangement is self equalizing. However, the power handling ability of each diode should be the same. In addition, the current ranges should be similar or the loads so arranged to avoid damaging any of the diodes.

A group of Zener diodes may be used as a voltage divider to obtain several regulated voltages simultaneously. (Fig. 11). This circuit may be used as a meter or 'scope calibrator. The four diodes shown in the circuit will supply 10 possible voltages:—

Voltage	Terminals
3.9	E – D
6.8	C – D
10.7	C – E
12.0	B – C
18.8	B – D
22.7	B – E
27.0	A – B
39.0	A – C
45.8	A – D
49.7	A – E

It may also be necessary at times to provide a regulated voltage lower than the 2.7 Volts minimum normally obtainable from a Zener diode. Voltages lower 2.7V may be obtained by using the *difference* in potential between a pair of Zeners, (Fig. 12). The temperature compensation of this circuit is excellent, for both Zener diodes tend to drift in the same direction, thus maintaining the difference voltage.

The second part of this series describes various ways in which Zener diodes are used in both ac and dc applications.

A PRACTICAL GUIDE TO ZENER DIODES

PART II

THE ZENER diode is generally associated with dc applications, such as the control and regulation of dc power supplies. Most Zeners are in fact used for this purpose but nevertheless they have many uses in ac, audio, rf, and ac control systems.

When supplied with alternating current and connected as a shunt regulator (Fig. 13), the Zener diode will limit both the positive and negative halves of the ac cycle. The diode conducts almost immediately after the signal passes through zero and into the negative segment. On the positive half cycle, the diode does not conduct until the applied voltage reaches the Zener voltage (Fig. 14). The result is a non-symmetrical square wave. Assymetry can be improved by using high input voltages, but can never be completely eliminated unless two shunt connected diodes are employed in a back-to-back configuration. (Fig. 15).

The Zener diode configuration shown in Fig. 15 is often used to provide stabilized filament supply voltages — especially to oscillator circuits and dc amplifiers. When using Zeners in this application bear in mind the ratio of average to peak Zener current. A figure of 0.6 is satisfactory.

Zener diodes may also be placed in the primary side of a step-down (or step-up transformer). When connected in this manner the diodes will regulate all associated secondary windings. The arrangement does require high voltage rated Zeners and is sometimes rather costly — however it is often used when high voltage secondary supplies need rudimentary stabilization.

Where power consumption is a prime consideration on ac power circuits, the Zener load resistor (RS) can be replaced by an inductance or capacitance. The device selected should have a reactance approximately the same as the calculated value for RS at the supply frequency.

OSCILLOSCOPE CALIBRATOR

A single Zener diode may be used as an inbuilt oscilloscope voltage calibrator that is independent of line voltage variations. Figure 16 shows how simply this facility may be incorporated in practically any oscilloscope. A selected 10 Volt Zener may be used to provide a calibration voltage of one volt per division.

ZENER NOISE VOLTAGES

As with neon regulator tubes, Zener diodes generate noise voltages. With Zener diodes, these voltages are associated with junction avalanche effects, and may vary between 10 uV and 1 mV depending upon the Zener type and voltage rating.

But unlike neon regulator tubes — where the incorporation of parallel filter capacitors is an excellent (if unintentional) way of making a relaxation oscillator — a Zener diode may be suppressed by adding parallel capacitance of 0.01 to 0.1 uF. This will reduce the noise voltage by a factor of at least 10 and yet maintain completely stable operation. (Figs. 17 & 18).

THE ZENER DIODE AS A FILTER

The Zener diode will respond to ripple voltages in much the same manner as it does with slow voltage variations. It has a very low dynamic impedance and thus reacts in much the same way as a filter capacitor.

Excellent power supply filtering can be obtained by connecting a Zener diode (having a Zener voltage equivalent to the ripple trough) across the load. In most circuit applications this will be as effective as adding a smoothing capacitor of several
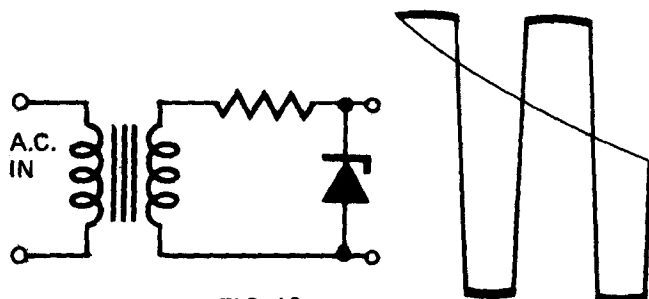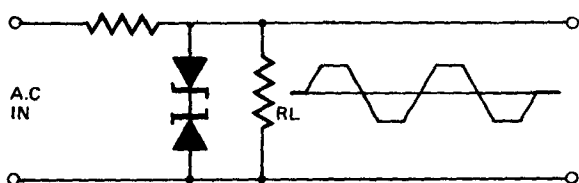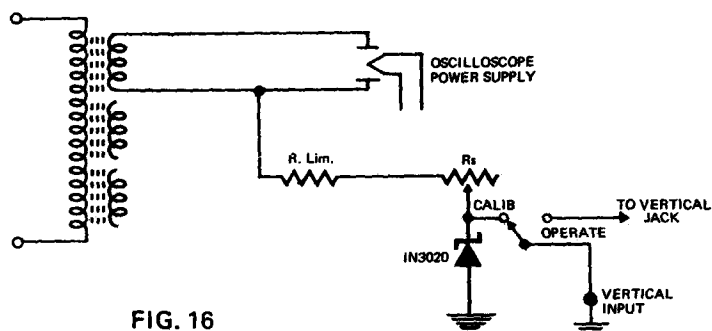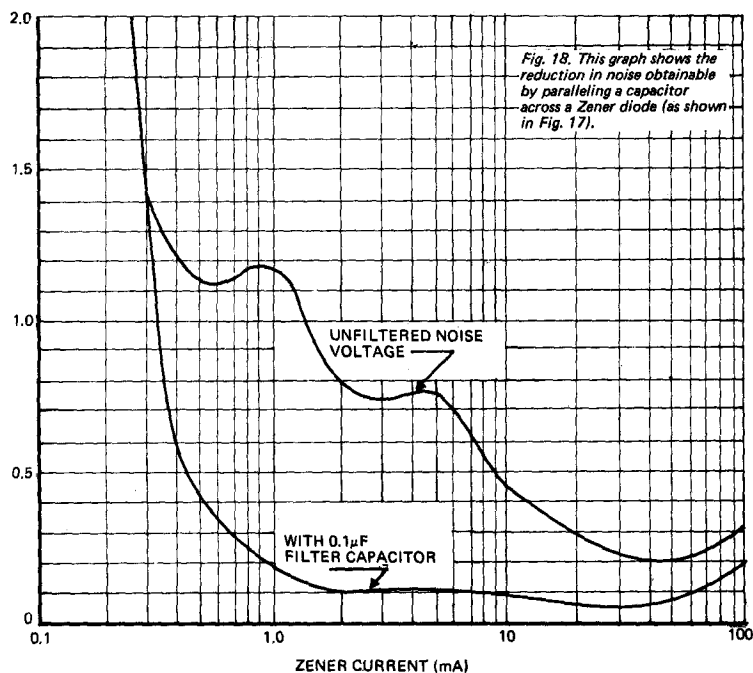

FIG. 13

FIG. 15

A.C. IN

FIG. 14

FIG. 16

Fig. 18. This graph shows the reduction in noise obtainable by paralleling a capacitor across a Zener diode (as shown in Fig. 17).
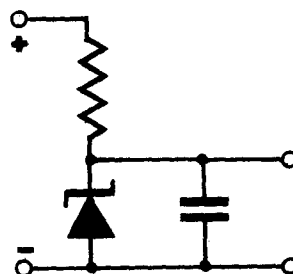
ZENER CURRENT (mA)

UNFILTERED NOISE VOLTAGE

WITH 0.1μF FILTER CAPACITOR

FIG. 17

FIG. 19

FIG. 20

FIG. 21

FIG. 22

thousand microfarads capacity, and will provide a considerable reduction in the level of ripple superimposed on the dc output.

INCREASING POWER HANDLING

Parallel connection is one way of increasing the power capabilities of Zener diodes. But a *simple* parallel connection (Fig. 19) is not practicable, for the Zeners will rarely be sufficiently well matched to conduct at exactly the same voltage. But by including very low resistance trimming resistors, the current levels can be matched so that both Zeners carry substantially equal amounts of the total current. (Fig. 20)

Although, as shown above, Zener diodes *can* be paralleled in order to increase their load carrying capacity it is usually more practicable to use a series or shunt transistor circuit with a Zener diode providing a voltage reference. This configuration will not only *improve the power handling capability by a factor of ten or so*, it will also improve the regulation of the circuit by an amount equal to the current gain of the transistor.

A simple Zener controlled shunt regulator is shown in Fig. 21.

The shunt regulator is very suitable for experimental and instructional use as it is totally short circuit proof. But, since maximum transistor current flows at zero load (and vice versa), it is also very inefficient, and for applications where power availability is limited the series transistor configuration is preferable.

A series regulator is shown in Fig. 22.

In this circuit the Zener diode establishes a reference voltage for the series transistor, which, in effect, operates as an emitter follower. Thus the emitter voltage is held within a few tenths of a volt of the base potential (which is determined by the Zener diode).

Thus the transistor acts as a series element to absorb voltage variations. All load current flows through this series transistor. The power handling ability of this type of supply is determined entirely by the number and type of transistors used (and the ability of the heat sink to remove heat). Figure 23 shows the regulation obtained from the simple circuit of Fig. 22 which uses a 4.7V ¼W Zener and a 1k series resistor. The regulation may be improved by a factor of 10 by substituting a low dynamic resistance 4.7V 3.5W Zener (we used an IR 1N1589) for the ¼W device.

This circuit (Fig. 22) can provide a variable voltage regulated output by connecting a 1k potentiometer across the Zener diode. The variable *reference voltage is then applied to the* base of the series transistor. However, this modification reduces the degree of regulation due to the shunting effect of the potentiometer. A better system is to switch different Zener diodes into the circuit for different voltage outputs.

CONSTANT CURRENT REGULATION

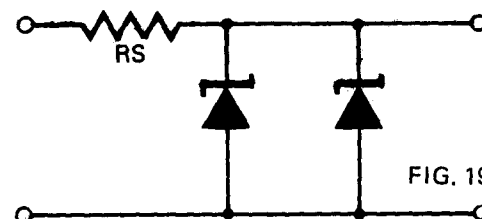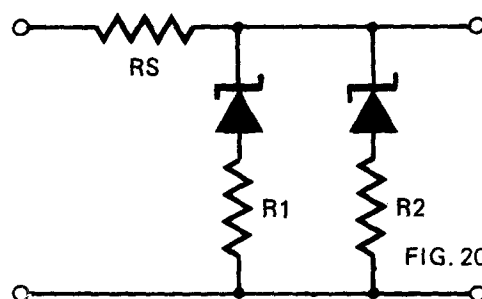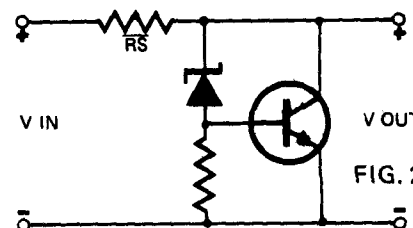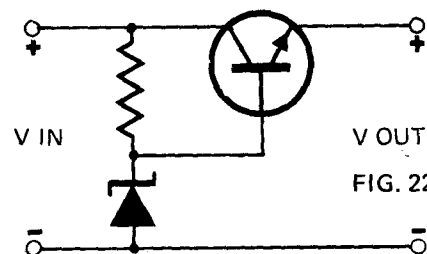A simple Zener-regulated constant current supply can be built using a single transistor as a variable series resistor. Figure 24 shows how it is done. Two circuit paths exist; one through the Zener diode which is in series with the bias resistor, and the other through R1, R2 and the 2N301 series transistor. Any change in the current through R3 causes a change in bias, the series transistor thus changes resistance to correct the current flow. In operation the current will remain substantially constant (within about 10%) from a short circuit to a 400 ohm output load. A graph of the output characteristics is shown in Fig. 25.

LOGIC CIRCUITS

In many ways a Zener diode

FIG. 23

FIG. 24
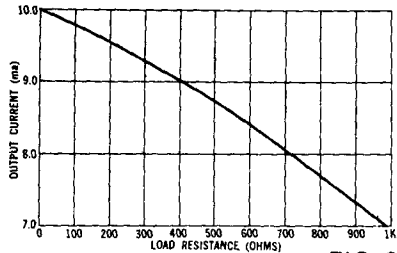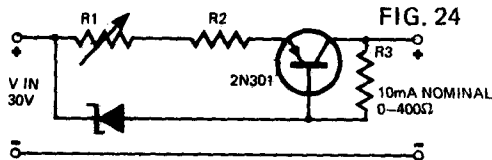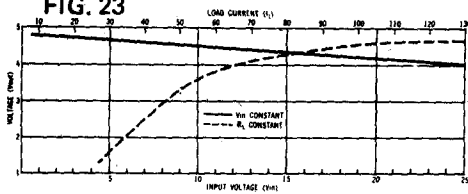
R1 R2

V IN
30V

R3
2N301
10mA NOMINAL
0–400Ω

FIG. 25

resembles a switch, and is therefore often used in computer and instrumentation logic circuitry.

The advantage of Zener diodes for this purpose is their extremely rapid operation when switching around their avalanche point. Whereas germanium or silicon diodes are limited to data rates of less than 2.5MHz (due to storage of minority carriers) Zener diodes switching about their avalanche point have switching times practically equal to their relaxation time. For silicon this is 10^{-9} seconds.

A typical AND gate, using two 6.8V Zener diodes is shown in Fig. 26. The positive 12 Volt bias is applied to both cathodes through the common load resistor, maintaining the diodes in avalanche condition. If a positive pulse is now applied to input A, the associated Zener will be gated out of the avalanche region. But due to the low impedance of the parallel connected Zener (associated with input B), the output (point A.B) remains clamped at 6.8Volts. The same circuit conditions apply if a pulse is applied solely to input B. However if a pulse arrives simultaneously at points A and B, both Zeners will be gated out of the avalanche region and the output voltage will rise from 6.8 Volts to the 12 Volt supply potential during the time that the pulse is coincident at the two inputs. Thus a positive pulse is produced coincident with the input pulse. Negative input pulses will not affect the circuit.

Figure 27 shows how a similar technique may be used to provide an OR function. In this configuration a pulse applied to *either* input will produce a pulse at the output.

SORTING

In the circuit shown in Fig. 28 the relays will be progressively energized as the input voltage increases. This circuit is often used for voltage controlled sorting. The relay contacts may be arranged to open chutes and illuminate indicator lamps for rapid sorting.

OVERVOLTAGE PROTECTION

The voltage sensitive characteristic of Zener diodes can be combined with the current sensitive characteristic of fuses to protect circuit components from overvoltage surges, whilst at the same time eliminating the 'nuisance' fusing that occurs when a fuse too close to the operating current is used.

By connecting a Zener diode of the correct voltage rating across the load, a fuse adequate to carry the normal load operating current for long periods may be used. But if the input voltage increases — and so exceeding the Zener breakdown voltage — the Zener diode will conduct. The sudden increase in current will blow the fuse practically instantaneously. (Fig. 29). A similar circuit may be used in conjunction with a circuit breaker rather than a fuse.

UNDERVOLTAGE PROTECTION

In some applications it may be necessary to disconnect a load from the mains supply if the supply voltage falls below critical level. A simple circuit that will provide this function is shown in Fig. 30. The series resistance R1 is chosen so that at normal operating voltages the Zener diode is broken down and sufficient current flows to hold the relay closed. When the supply voltage falls below the desired level, the Zener ceases to conduct and the relay drops out. The addition of the Zener diode to this circuit proivides an accurate reference point, increasing reliability and eliminating the need for specially selected relays for different voltages.

DUAL VOLTAGE SUPPLY

Most logic circuitry needs a dual power supply — (one positive and one negative with respect to zero). The useful, but little known circuit shown in Fig. 31 can supply a dual output of balanced or unbalanced voltage from a single ended power supply. Zener diodes should be chosen to suit the voltages required.

Full details of operating characteristics of Zener diodes can be obtained from most semiconductor manufacturers. ●
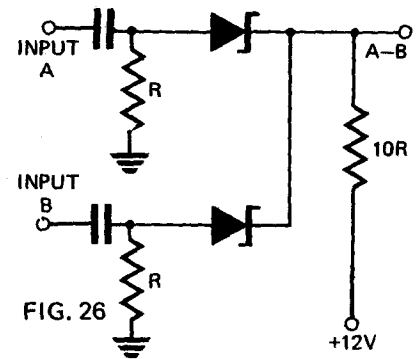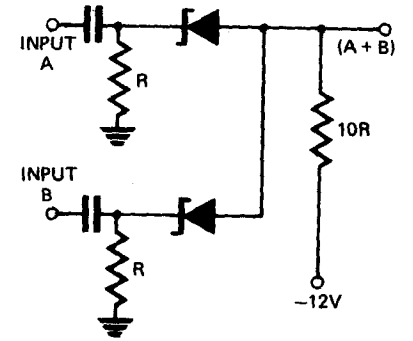
INPUT A
INPUT B
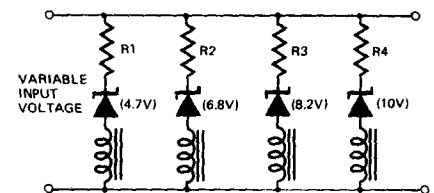R
R
A–B
10R
FIG. 26
+12V

INPUT A
INPUT B
R
R
R
(A + B)
10R
–12V
FIG. 27

VARIABLE INPUT VOLTAGE
R1 R2 R3 R4
(4.7V) (6.8V) (8.2V) (10V)
FIG. 28
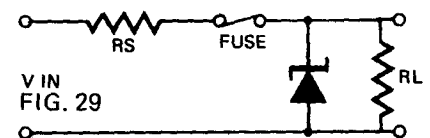
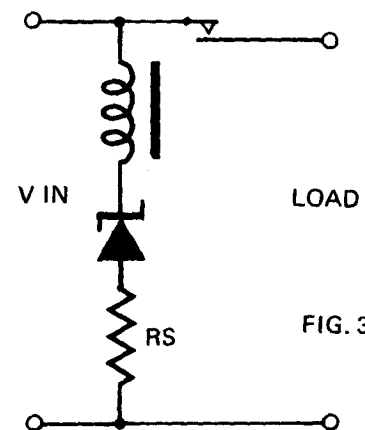RS FUSE RL
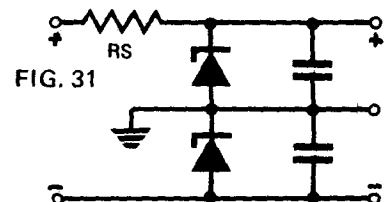V IN
FIG. 29

V IN LOAD
RS
FIG. 30

RS
FIG. 31

MODERN CRYSTAL OSCILLATORS

Described by Roger Harrison VK2ZTB

CRYSTAL OSCILLATORS, in one form or another, are fundamentally associated with virtually all transmitting and receiving equipment. Basic circuitry and circuit techniques, and the fundamentals of quartz crystals are discussed at various length by both the ARRL and RSGB handbooks, Pat Hawker's 'Amateur Radio Techniques', the various VHF handbooks by Jessop (RSGB) and Tilton (ARRL) as well as 'The Radio Handbook' by Bill Orr (Editors and Engineers). A useful, and more recent discussion on the subject of crystals and crystal oscillators is contained in the 'Ham Notebook' from the editors of the American journal 'Ham Radio'. For a deeper appreciation of the subject, references (1) to (4) are recommended.

Basic solid state crystal oscillator circuit techniques are by now well established, most circuits being adaptations of the well-known vacuum tube technology such as the Pierce, Hartley, Clapp and Butler oscillator and use both bipolar and FET devices. Whilst these circuits basically fulfil their intended purpose, there are many applications which require something different or where performance needs to be reliably characterised.

Presented here are a variety of circuits, for a range of applications from LF through the VHF range, that are not commonly found in current amateur use or literature.

MODES OF OPERATION

A point not often appreciated, or just forgotten, is that quartz crystals can oscillate in a *parallel* resonant mode *and* a *series* resonant mode. The two frequencies are separated by a small amount, typically 2-15 kHz over the frequency range. The series resonant frequency is *lower* in frequency than the parallel. A crystal specified and calibrated for use in the parallel mode may be satisfactorily used in a series resonant circuit if a capacitor equal in value to its specified load capacitance (usually 20,30, 50 or 100 pF) is connected in series with the crystal. Sadly, you can't invert the process for series resonant crystal in parallel mode circuits. The series mode crystal will oscillate higher than its calibrated frequency in this case and it may not be possible to capacitively load it down sufficiently.

Overtone crystals operate in the *series* mode usually on the third, fifth or seventh overtone, and the manufacturer normally calibrates the crystal *at the overtone frequency.* Operating a crystal in the parallel mode and multiplying the frequency three or five times produces quite a different result from operating the same crystal in the series mode on its third or fifth overtone. When ordering overtone crystals avoid confusion and specify the frequency you want, *not* the apparent fundamental frequency. Reference (4) makes this point quite clear.

Fundamental crystals in the range 500 kHz to 20 MHz are usually specified for parallel mode operation but series mode operation can be requested. For low frequency crystals, up to 1 MHz, either mode can be specified. Overtone crystals generally cover the range 15 MHz to 150 MHz.

WIDE RANGE or A PERIODIC OSCILLATORS

Oscillators that do not employ tuned circuits can be very useful, whether they are simply used as 'crystal checkers' or some other purpose. Particularly for LF crystals, tuned circuits can be bulky. However, they aren't without their traps. Some crystals are prone to oscillation on unwanted modes, particularly the DT and CT cut crystals used for LF quartz oscillators. It is wise to check that the output is on the correct frequency and no mode instability is evident. Reducing feedback at the higher frequencies usually cures this. In extreme cases, the idea has to be abandoned and an oscillator having a tuned circuit used instead, (LF crystal oscillators are discussed later).

The first circuit is an emitter-coupled oscillator, a version of the Butler circuit. The basic circuit first appeared in VHF Communications in 1970 (p.240) as portion of a VHF-UHF calibration spectrum generator. Versions have subsequently been published in the 'VK5 Bulletin' (S.A. Div. WIA) in 1972 and 6UP, August issue, 1974. Lane (3) discusses a variation of this circuit (Fig. 2).
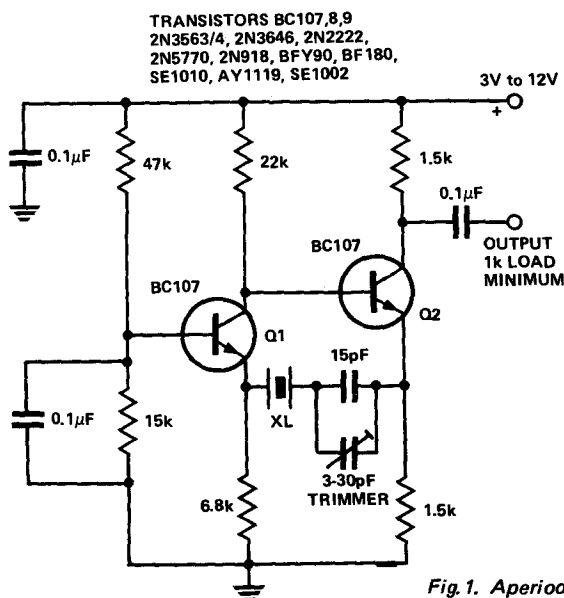
The output of the circuit in Fig. 1 is

TRANSISTORS BC107,8,9
2N3563/4, 2N3646, 2N2222,
2N5770, 2N918, BFY90, BF180,
SE1010, AY1119, SE1002
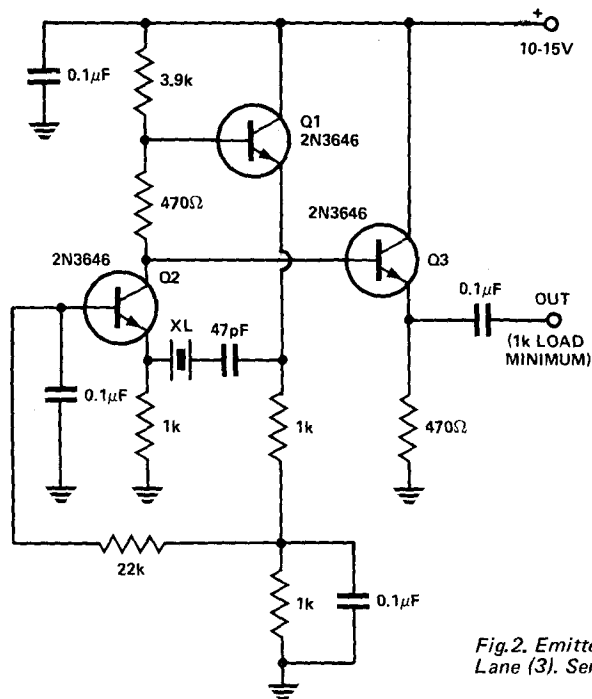
Fig. 1. Aperiodic Butler oscillator (series mode)

Fig.2. Emitter-coupled oscillator — after Lane (3). Series mode.

essentially sine wave; reducing the emitter resistor of Q2 increases the harmonic output. By doing this, a 100 kHz crystal produces good harmonics through 30 MHz. It is a series mode circuit.

A variety of transistors may be used. For crystals above 3 MHz, transistors with a high gain-bandwidth product are recommended. For crystals in the 50 kHz to 500 kHz range, transistors with high LF gain, such as the 2N3565 are recommended. Also, for crystals in this range, permissable dissipation is usually less than 100 microwatts and amplitude limiting may be necessary. Low supply voltage, consistent with reliable starting, is recommended. Modifying the circuit by the addition of diodes — as shown in Fig. 3 — is a better method, and starting performance is improved. The circuit will oscillate up to at least 10 MHz with appropriate transistors and emitter resistor

values. An emitter follower or source follower buffer is recommended. Similar comments to the above apply to Fig. 2. An emitter follower buffer is included in this circuit. Both circuits are slightly frequency sensitive to power supply voltage changes and load variations. A load of 1 k or greater is recommended.

TTL IC can be used in crystal oscillator circuits but many published circuits have poor starting performance or suffer from non-repeatability owing to wide paramater spreads in IC's. The circuit in Fig. 4. is by K1PLP from QST, Feb. 1974 (5) and is after Weggeman (6). This circuit has been tried by the writer over the range 1 MHz to 18 MHz and can be recommended. It is a series mode oscillator and suits AT-cut crystals. The output is about 3 volts peak to peak, square wave up to about 5 MHz beyond which it is becomes

more like half-sine pulses. Starting performance is excellent, often a critical factor with TTL oscillators.

LOW FREQUENCY CRYSTAL OSCILLATORS

Crystals in the range 50 kHz to 500 kHz require special considerations not encountered with the more common AT or BT cut HF crystals. The equivalent series resistance (which determines 'activity' — that figure of merit of days of old) is much greater and their permissable dissipation is limited to less than 100 microwatts, preferably 50 microwatts or less.

The circuit in Fig. 5. is a series mode oscillator described by Lane (3). It has the advantage of ot requiring a tuned circuit, and has a choice of sine or square wave output. For crystals in the range 50-150 kHz, 2N3565 transistors are recommended although the author has found BC107's satisfactory. Either type will suffice for crystals in the range 150 kHz to 500 kHz. If you find the crystal will not start reliably, most likely the crystal has a very high equivalent series resistance, in which case increase R1 to 270 ohms and R2 to 3.3 k (as recommended by Lane). For square wave operation, C1 is 1 uF (or a value close to, or above it). For sine wave output, C1 is not in circuit. Amplitude limiting is unnecessary. Sine wave output is about 1 V rms, square wave output about 4 V peak to peak.

The circuit in Fig. 6 is also described by Lane (3) and can be recognised as a modified form of the Colpitts oscillator, with the addition of resistor Rf to control feedback (it works the same way as Eno's). Capacitors C1 and C2 should be reduced by preferred values as the frequency is increased. At 500 kHz, values for C1 and C2 should be around 100 pF and 1500 pF respectively.

The circuit as shown gives sine wave output with the second harmonic
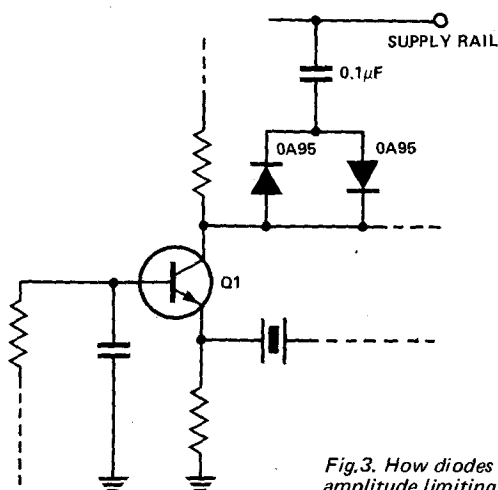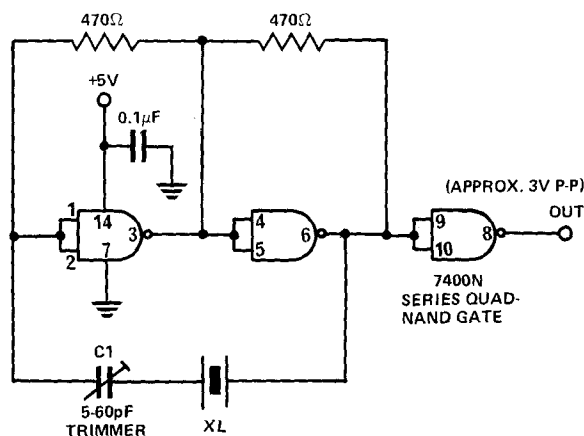

Fig.3. How diodes are used for amplitude limiting.

Fig.4. Reliable TTL crystal oscillator.
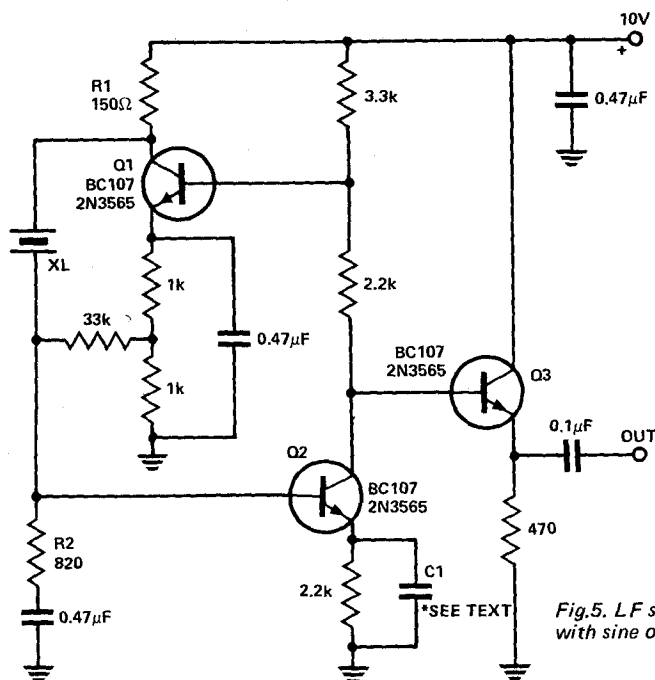
MODERN CRYSTAL OSCILLATORS

10V

R1 150Ω
Q1 BC107 2N3565
XL
3.3k
0.47µF
1k
33k
0.47µF
2.2k
1k
BC107 2N3565 Q3
0.1µF OUT
Q2
BC107 2N3565
470
R2 820
2.2k
C1 *SEE TEXT
0.47µF

Fig.5. LF series mode oscillator with sine or squarewave output.

12V

Rf 3.3k
120k
2.2k
0.1µF
XL
5-60pF
BC107
OUTPUT (MIN .1k LOAD)
100pF
C2
C1 .0033pF
.0039µF
10k
100Ω
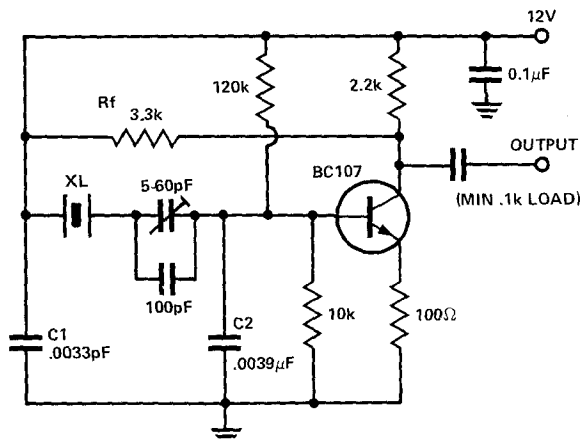
Fig.6. Parallel mode LF oscillator.

about 40 dB down (or greater). This can be reduced by careful trimming of Rf and C1. Note that, at the reduced level of feedback necessary to achieve this, it takes some 20 seconds for the oscillator to reach full output. Output is about 2 to 3 volts peak to peak.

If you need an output rich in harmonics, the simple addition of a 0.1 uF capacitor across the emitter resistor will achieve this. Output then rises to about 5 V peak to peak. Power supply voltage can be reduced in this case to lower crystal dissipation.

Other transistors can be used, but bias and feedback may have to be adjusted. For cantankerous crystals determined to oscillate in modes other than those you wish, the circuit of Fig.7 is recommended. Feedback is controlled by tapping down the collector load of Q1. Amplitude limiting is necessary to keep the crystal dissipation within limits. For 50 kHz crystals the coil should be 2 mH and its resonating capacitor 0.01 uF. Output is about 0.5 V rms, essentially sine wave. The use of an emitter follower or source follower buffer is recommended. If a parallel mode crystal is used the 1000 pF capacitor shown in series with the crystal should be changed to the crystal's specified load capacitance (usually 30, 50 to 100 pF for these crystals).

HF CRYSTAL OSCILLATOR CIRCUITS

Solid state circuits for the popular AT-cut HF crystals are legion. However, results aren't always what one would expect. Most fundamental crystals up to 20 MHz are ,usually specified for parallel mode operation. However, such crystals can be used in series mode oscillators by putting the specified load capacitance in series with the crystal as mentioned previously. Both types of circuit are detailed here.

A useful oscillator for the range 3 to 10 MHz that does not require a tuned circuit is given in Fig. 8 (a). It is, of course, the same circuit as Fig.6. The circuit can be used down to 1 MHz if C1 and C2 are increased to 470 pF and 820 pF respectively. It can be used up to 15 MHz if C1 and C2 are reduced to 120 pF and 330 pF. Respectively. This circuit is recommended for non-critical applications where high harmonic output is wanted, or not a consideration.

The addition of a tuned circuit as in 8(b) reduces harmonic output considerably. A tuned circuit with as high a Q as possible is recommended. In a 6 MHz oscillator, I have obtained the following results. With a coil Q of 50 the 2nd harmonic was 35 dB down. With a Q of 160, it was −50 dB! Resistor Rf can be adjusted (increase slightly) to improve this. The output is also increased with a high Q coil. As previously noted, with reduced feedback it takes some tens of seconds to each full output from switch on, however, frequency stability is excellent.

Operation at other frequencies is accomplished by changing the capacitors and coil appropriately.

This circuit (Fig. 8) can also be turned into a very effective VXO. A small inductance is placed in series
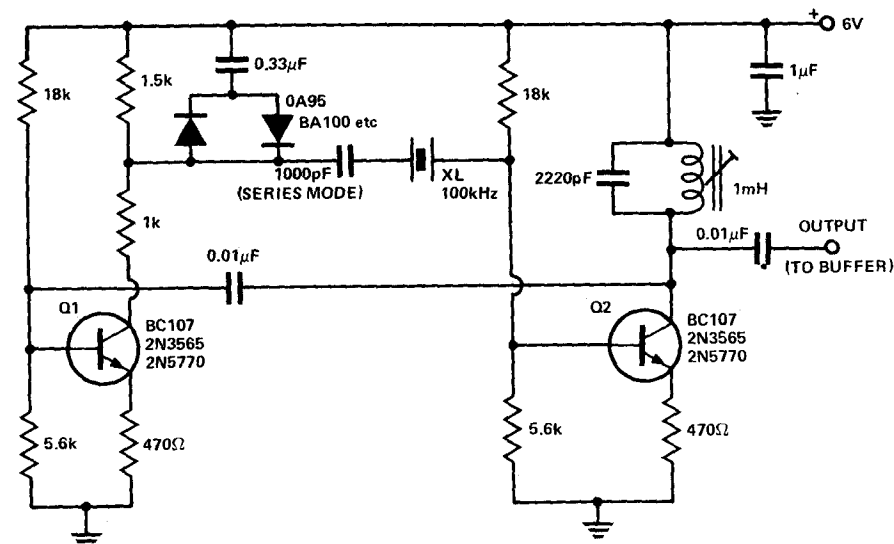
6V

18k
1.5k
0.33µF
0A95 BA100 etc
18k
1µF
1000pF (SERIES MODE)
XL 100kHz
2220pF
1mH
1k
0.01µF
OUTPUT (TO BUFFER)
Q1 BC107 2N3565 2N5770
Q2 BC107 2N3565 2N5770
5.6k
470Ω
5.6k
470Ω

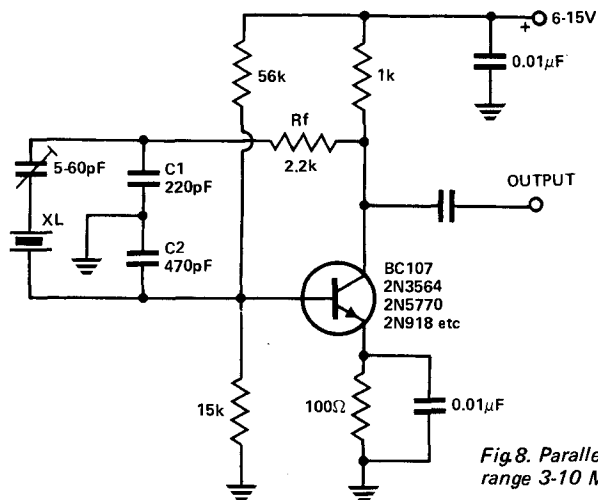Fig.7. 100 kHz crystal oscillator (with tuned circuit).
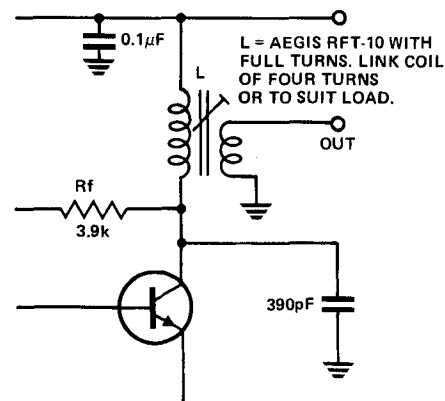
Fig.8. Parallel oscillator for the range 3-10 MHz.

Fig.8b. Adding a coil to the circuit shown in Fig.8.

with the crystal and one of the capacitors in the feedback circuit is made variable. An ordinary two-gang 10-415 pF (or thereabouts) broadcast tuning capacitor will do the job nicely. Both gangs are paralleled. The tuning range depends on the crystal used, the inductance of L1 and the frequency. A greater range is usually obtainable with the higher frequency crystals. Stability is excellent, approaching that of the crystal.

Another variation of this circuit is shown in Fig. 10. This circuit may allow more 'pull' on the crystal, but stability is poorer. For both Fig's 9 and 10 the trimmer is to set the nominal frequency at some position of the tuning capacitor. For both circuits also, especially for Fig.10, the output varies across the tuning range.

A VHF OSCILLATOR-MULTIPLIER

The circuit in Fig.11 is a modification of the 'Impedance Inverting' overtone oscillator discussed by Rankin (4), who also describes a similar circuit (albeit outdated — using an OC171 — even Dick Smith no longer stocks them!) Normally, with the impedance inverting circuit, the collector is either untuned or grounded for RF. The collector can be tuned to twice or three times the crystal frequency To reduce the output at the crystal frequency, a double tuned circuit is recommended. DO NOT tune the collector to the crystal frequency, otherwise the circuit will oscillate at a frequency not controlled by the crystal. It is advisable to keep the collector lead as short and direct as possible.

Results with this circuit are excellent. All outputs other than the wanted output were at −60 dB or greater. Noise output is at least 70 dB below the wanted output. It makes an excellent conversion oscillator for VHF/UHF converters. Almost 2 V of RF is available at the hot end of L3
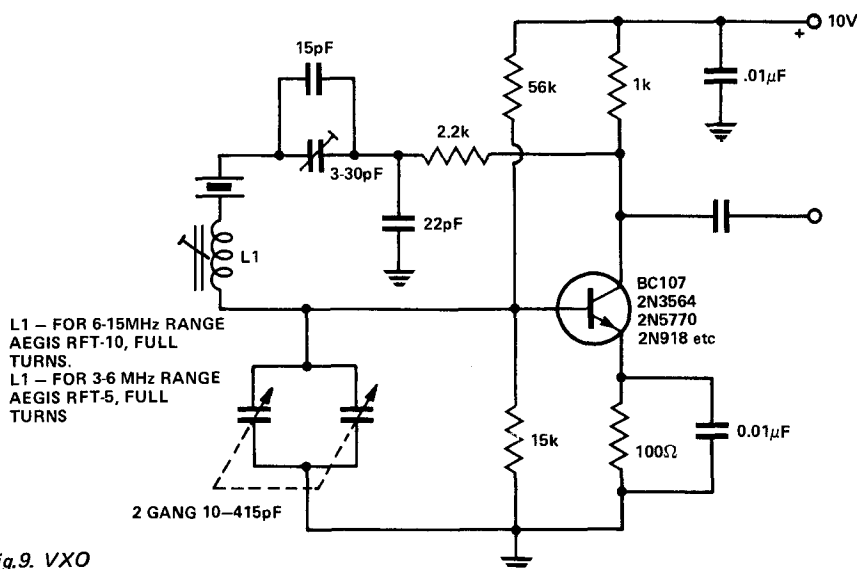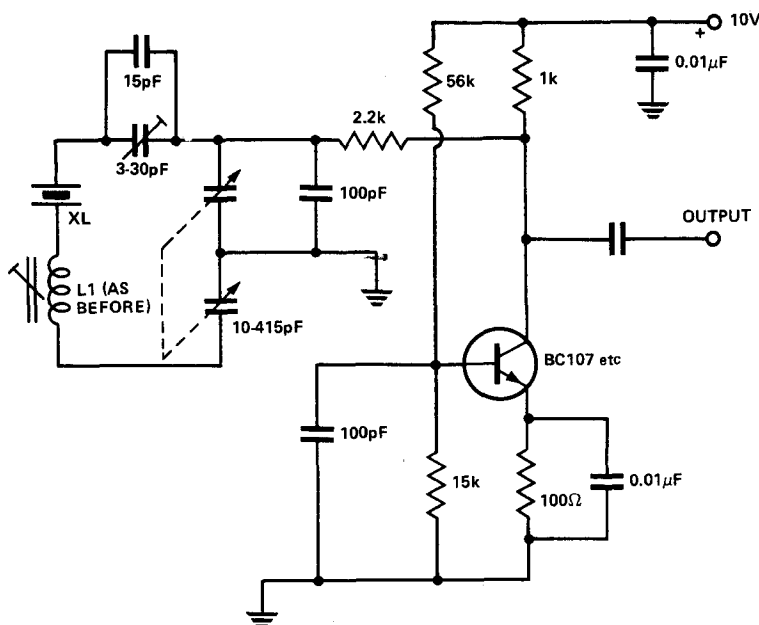

Fig.9. VXO

Fig.10. Alternative version of circuit shown in Fig.9.

145

MODERN CRYSTAL OSCILLATORS

(author's prototype at 30 MHz). A Zener regulated supply is recommended. As indicated on the diagram, different circuit values are necessary for different transistors. Strays in individual construction may also necessitate variations. L1 can be used to pull the crystal onto frequency.

Slight variations in frequency (about 1 ppm) occur when tuning L2 and L3 and also with load variations. However, in practise, these turn out to be of no consequence.

REFERENCES

(1) 'Radio Transmitters', L. Gray & R. Graham (McGraw-Hill)

(2) 'Electronic Fundamentals & Applications', J. D. Ryder (Pitman)

(3) 'Transistor Crystal Oscillators to Cover Frequency Range from 1 kHz to 100 MHz' by M. Lane, Australian Post Office Research Laboratories, Report No. 6513.

(4) 'Overtone Operation of Quartz Crystals' D. Rankin (VK3QV), Amateur Radio, March and May 1967.

(5) 'A TTL Crystal Oscillator', K1PLP, QST February 1974, p.34.

(6) 'IC-Compatible Crystal Oscillator', The Electronic Engineer, May 1969.

Fig. 11. VHF oscillator — multiplier.
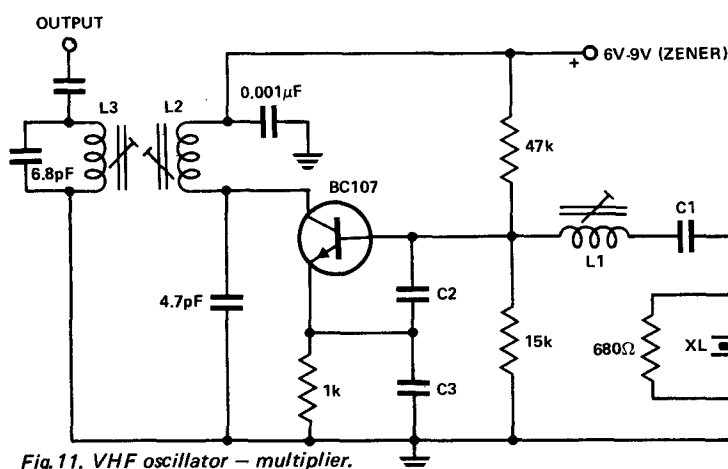
65 MHz Xtal 130 MHz OUTPUT
L1 = NEOSID AZ ASSEMBLY (4mm FORMER & F29 SLUG) WOUND WITH 12 TURNS OF 4.55 mm ENAMEL WIRE, CLOSEWOUND

L2/3 = NEOSID, DOUBLE ASSEMBLY 7300 CAN, TWO 722/1 FORMERS, F29 SLUGS, WOUND WITH 5 TURNS, 0.63mm ENAMEL, CLOSEWOUND

C1 = 33 OR 39 pF

43 MHz Xtal 130 MHz OUTPUT
L1 = 20 TURNS 0.4mm AS ABOVE
L2/3 = AS ABOVE
C1 = 56pF

38MHz Xtal 116 MHz OUTPUT
L1 = 24 TURNS 0.4mm AS ABOVE
L2/3 = 9 TURNS 0.63mm AS ABOVE
C1 = 68pF OR 100pF

XL	C2	C3	
65mHz	8.2pF	5.6pF	BC107
43mHz	15pF	10pF	BC107
38mHz	22pF	18pF	
65mHz	18pF	12pF	2N3564
43mHz	33pF	18pF	2N3564 / 2N5770
38mHz	56pF	39pF	2N5770